PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES
Volume 21 July 15, 1935 Number 7

THE TEMPO OF GROWTH OF FRATERNITIES

By Franz Boas

Department of Anthropology, Columbia University

Read before the Academy, Tuesday, April 23, 1935

On a previous occasion[1] I have discussed the results of measurements of children observed during consecutive years and covering a period of about 12 years, from 6 years to 17 years of age. Only the results of the observations on stature and weight are available. The measurements show the typical decrease in the rapidity of growth during childhood, the sudden spurt during adolescence and the rapid decrease of rate of growth after sexual maturity has been attained. This is the gross picture. The individual curves of growth during the period of adolescence show that the time of maximum rate of growth and that of puberty in girls are not in a fixed relation, although the maximum rate of growth precedes completed puberty. I proceeded to classify the growth curves of both boys and girls by years in such a way that all those who had their maximum rate of growth at the same age were grouped together. The girls were also so grouped that those who had their first menstruation at the same age were combined. These studies proved that when the period of maximum rate of growth is early, the whole growth proceeds at a rapid rate from early years on, at least as far as our material permits us to follow it; the onset of the puberty spurt is early, its intensity great, its duration short. Conversely, when the period of maximum rate of growth is late its intensity is slight, its duration long. This is expressed by the values of the maximum rates of growth which are low for those who reach maturity late, high for those who reach it early.

It is a most important question to decide whether this unity of the tempo of development that prevails until maturity has been reached will extend over later life. As far as I am aware the only reliable material that shows an interrelation of phenomena of aging and of the life span is that discovered by Dr. Bernstein, who has proved that early presbyopia indicates an early death by senile degeneration. All others are more or less impressions of medical practitioners who believe, for instance, that early calcification of the larynx indicates an early onset of arterial de-

generation. It seems most important that by organized effort the life history of individuals should be followed up in order to show whether the rate at which physiological development and decay occur are constitutionally determined. Studies on longevity also suggest that this is the case, but they can never be quite convincing.

The observations here discussed refer to individual development. We had to investigate next whether the conditions that determine the speed of the life cycle are hereditary or determined by environment. If they are hereditary we may expect that members of fraternities are alike in

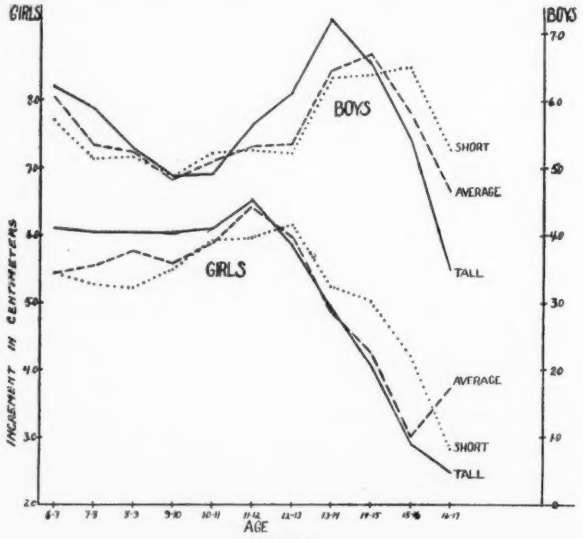


FIGURE 1

Growth of brothers and sisters according to stature of brother or sister at different ages (7, 9, 11 and 13 years) Hebrew Orphan Asylum siblings.

their tempo, provided the social conditions of the whole community investigated are fairly uniform. The best material at my disposal is contained in the records of the Hebrew Orphan Asylum in New York, for here the conditions of nutrition, shelter and mode of life are as uniform as they can be obtained. I have already shown that there is a fairly strong positive correlation between the dates of first menstruation of sisters.[2] Unfortunately it was not possible to classify the material relating to growth in the same manner as was done for individuals because, on account of gaps in the series of observations, the moment of maximum rate of growth cannot be determined for all the members of each family

and the material would have been too scanty. On account of the asymetries of the increment curve, particularly during adolescence, I did not use correlations, but classified the material in three groups, about equal in numbers: children tall, medium and short at a given age and compared the curves of growth of their brothers and sisters. Then it appeared that the brothers and sisters of the tall ones, who include many of those with rapid tempo of development, will also have a rapid tempo, an early

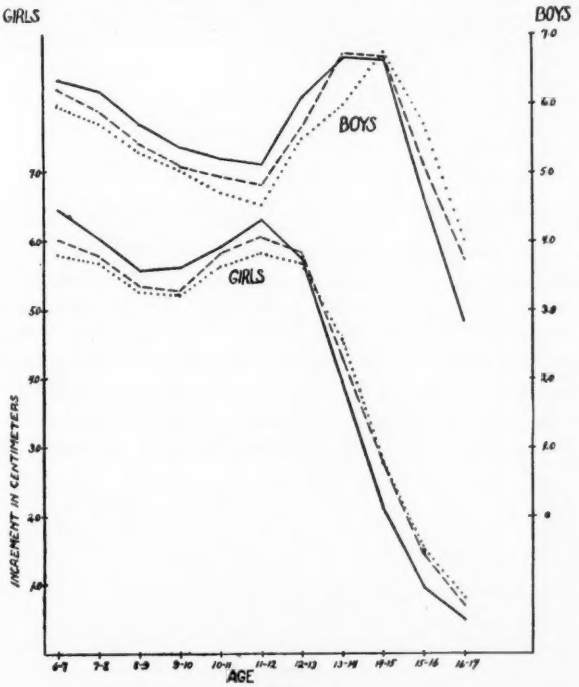


FIGURE 2

Growth of brothers and sisters according to stature of brother or sister at different ages (7, 9, 11 and 13 years) Horace Mann School siblings.

time for the maximum rate of growth, a rapid rate and an early termination of growth, while the brothers and sisters of the short ones, who include many of those with sluggish tempo of development, have a slow rate of growth of less intensity and longer duration (Fig. 1). Since the conditions under which these children live are unusually uniform, we may conclude that proof for the heredity of the tempo of growth has been given. This agrees with the results obtained by Pearl by experimentation with animals

and with observations on plants of short and long vegetative periods. An analogous study of children in the Horace Mann School gave the same result. In this case it might be suspected that similar home environment is a contributory cause to the similarity of the growth curve of members of each fraternity (Fig. 2).

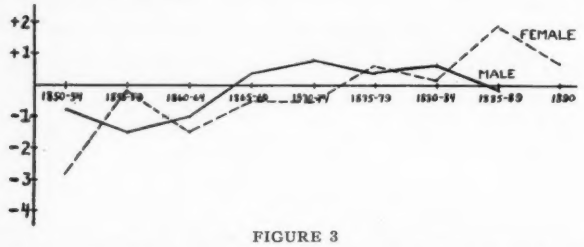


FIGURE 3

Deviations according to year of birth from mean statures of immigrants.

It is obvious that a phenomenon of such complexity as length of body and tempo of development must be governed by many hereditary factors and that we are dealing with a phenomenon of general organization of the body and that a search for genes would not be advisable. Is not there some danger anyway, that the number of genes will depend rather upon the number of investigators than upon their actual existence?

It is well known that stature is strongly influenced by peristatic conditions. Not only has there been a constant increase in stature during the

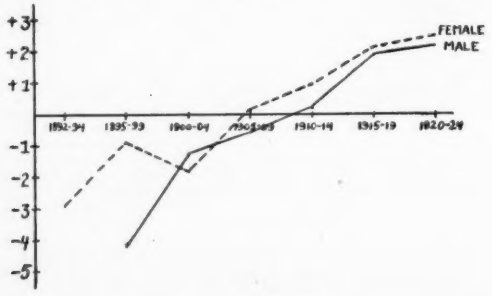


FIGURE 4

Deviations according to year of birth from mean statures of children in Hebrew Orphan Asylum, New York City.

past century, but various studies have also proved that children of parents living under modern conditions exceed in stature their parents. The recent study of the stature of Harvard students compared with those of their fathers[3] demonstrates this as well as the earlier studies of the statures of immigrants and of their children.[4] I have investigated the statures of

Jewish immigrants who came to the United States who were measured in 1910, and were born during the period from 1850 to 1890. These do not show the characteristic increase of stature of European populations but merely indicate the slow loss of stature of the individual from 35 years on (Fig. 3). If there is any increase it is very slight. This would indicate that the conditions of growth of this particular part of the European population have not changed. In contrast to this the statures of growing children of these same immigrants, living under New York conditions, demonstrate a constant increase according to the year of their

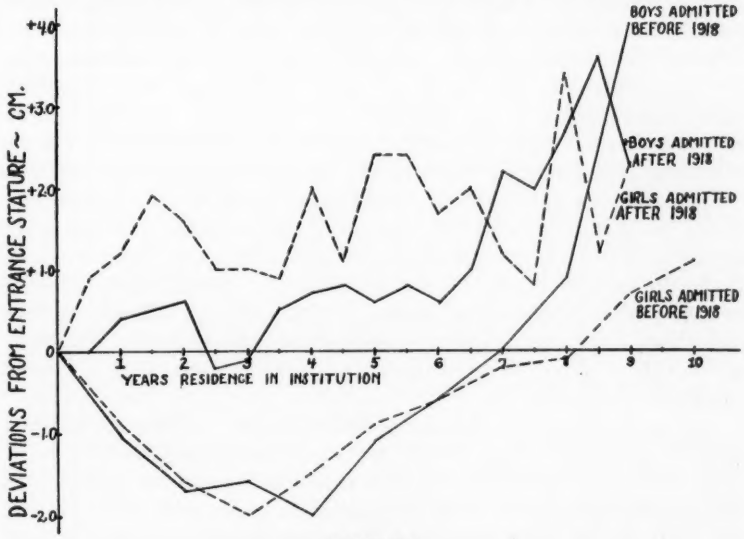


FIGURE 5

Deviations from average entrance stature after given period of residence of children in Hebrew Orphan Asylum, New York City.

birth (Fig. 4). The data cover the period from 1892 to 1924 and during this time the children exceed the averages of their ages the more the later they are born. There are no clear indications of irregularities in this curve except perhaps during the years of the war. This is the more remarkable since these children were growing up in the poorest parts of the city.

The study also gave an opportunity to inquire into the influence of institutional life upon the development of children. Up to the year 1918 the management of the Hebrew Orphan Asylum was rigid. The children were fed adequately in quantity, but the food was not well balanced. There was strict discipline and little opportunity for free play. When

comparing the children entering during this time with those who had been subjected to the institutional environment a decided retarding effect made itself felt (Fig. 5). For a number of years the children were in stature the more inferior to those entering, the longer they had been there. Many stayed until after 1918, when the management of the Institution was radically changed. This is expressed by a gradual rise of the stature of the children who had been in the Institution for a long time until finally they reached normal size.

By contrast the children admitted after 1918 developed much better. At that time the food supply was properly regulated, medical supervision became much stricter and care was taken to give the children fresh air, freedom and bodily exercise. The effect upon growth was immediate and we see that the children are the taller for their age, the longer they have been subjected to the régime of the Institution. It may be considered as proved that it is not the institutional life that hinders the development of children, but rather the way in which the institutions are run.

Stature and weight alone are inadequate measures of development. I may recall here studies in the Hebrew Home for Infants of New York in which on the whole, children of very poor hereditary stock are found. Under the excellent physical care that they enjoy, these children are larger in bulk, that is, in stature and weight, than the average American children of the same ages, according to the large sample collected by Dr. Woodbury all over the country; but their dentition lags far behind that of normal children.[5]

The importance of environmental influences may be observed in other bodily traits. I have been able to show by an extended series of measurements that the cephalic index is not stable when families are transplanted into a new environment. The measurements which were taken in 1909 and 1910 gave for immigrant East European Hebrews an index of 83.0, for their American-born descendants an index of 81.4. At that time Hebrew families with native-born parents were rare. A selection of those with American-born mothers gave an index of 79.7. At present it is easier to find such families and an adequate number is being studied by Dr. Nicholas Michelson. He finds from a preliminary tabulation an index of 78.7 for the descendants of American-born East European Hebrews.

All this shows that our attention must be directed equally to the hereditary and environmental influences determining bodily form.

[1] *Human Biology*, **4**, p. 307 et seq. (1932); **5**, p. 429 et seq. (1933).

[2] *Ibid.*, **4**, p. 308.

[3] G. J. Bowles, *New Types of Old Americans at Harvard*, 1932.

[4] F. Boas, *Changes in Bodily Form of Descendants of Immigrants*, 1912.

[5] F. Boas, "Eruption of Deciduous Teeth Among Hebrew Infants," *Jour. Dent. Res.*, **7**, 245 et seq.

ANALYSIS OF 18,000 PROPER MOTIONS DERIVED AT THE LEANDER McCORMICK OBSERVATORY

BY P. VAN DE KAMP AND A. N. VYSSOTSKY

LEANDER McCORMICK OBSERVATORY, UNIVERSITY OF VIRGINIA

Read before the Academy, Monday, April 22, 1935

1. *Introduction.*—The relative proper motions of 18,000 faint stars photographed with the 26-inch visual refractor of the Leander McCormick Observatory were measured in 341 regions fairly uniformly distributed north of $-30°$ declination. Each region contains at least one star with known absolute proper motion which served to reduce the relative proper motions to absolute and which was cut down by the rotating sector to the tenth magnitude. Since most of the reduction stars are in Boss's Preliminary General Catalogue we shall call them "Boss" stars, although for the sake of greater uniformity in distribution regions were added containing a star with the absolute motion newly derived at the Dudley Observatory. The absolute proper motions of 574 stars were kindly furnished by the Dudley Observatory in advance of publication of the New General Catalogue. The system of proper motions is that of the P. G. C. with Raymond's correction $\Delta\delta_s$ applied to the motions in declination.

The probable error of a relative motion amounts to about $0''.005$, the probable error of the reduction from relative to absolute motion was found to be about $0''.007$ per Boss star, or since there are on the average 1.7 Boss stars per region, $0''.005$ per region.

Photovisual magnitudes were derived by a direct comparison with the North Polar Sequence; the average limiting magnitude was found to be 12.5 pv. For 5200 stars, spectra were secured at the Harvard College Observatory.

2. *General Solution for Solar Motion, Differential Galactic Rotation and Precessional Corrections.*—The absolute proper motions furnished by the Dudley Observatory were derived with Newcomb's precession uncorrected for the motion of the equinox. It was decided to apply as precessional corrections:

$$\Delta\mu_\alpha \cos\delta = +0''.0015 \cos\delta - 0''.0041 \sin\alpha \sin\delta$$
$$\Delta\mu_\delta = -0''.0041 \cos\alpha$$

These corrections are based on the mean of two determinations by Oort,[1] and by Plaskett and Pearce:[2]

$$-\Delta e - \Delta\lambda = -0''.0110 \quad \text{and} \quad \Delta p = +0''.0104$$

where Δe, $\Delta\lambda$ and Δp represent corrections to Newcomb's values of the motion of the equinox, planetary and luni-solar precession.

After the relative proper motions were reduced to absolute, they were transformed to the galactic system of coördinates, using van Rhijn's galactic pole[3] ($\alpha = 12^h56^m$, $\delta = +25^\circ.5$). For all stars of magnitude 7.5 and fainter mean values of $\mu_l \cos b$ and μ_b were formed in each region and used in a simultaneous analysis for precessional corrections, differential galactic rotation and solar motion. In order to reduce the number of equations, the 341 regions were combined into 105 groups. Each group contains on the average 3.2 regions, 5.5 Boss stars and 170 faint stars. Thus we had 210 equations of condition of the usual form, with eight un-

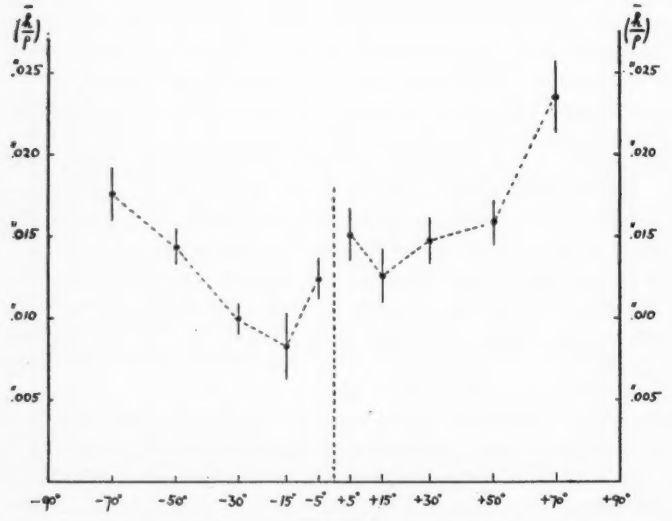


FIGURE 1

Mean secular parallax for different galactic zones (all stars, mean magnitude 11.3). The half-lengths of the vertical lines through each point indicate probable errors.

knowns, namely, the components of solar motion X, Y, Z, the constants of differential galactic rotation Q, $C = P \cos 2l_o$, $S = P \sin 2l_o$ and the corrections $\Delta' p$ and $-\Delta' e - \Delta' \lambda$ to the precessional corrections already applied. Here l_o is the longitude of the axis of rotation, $Q = \frac{B}{4.74}$, $P = \frac{A}{4.74}$; A and B are Oort's constants.

Before embarking on the general solution it was necessary to make a provisional solution for solar motion for different galactic latitude zones

in order to find the dependence of secular parallax $\left(\frac{\bar{h}}{\rho}\right)$ and apex (longitude l_a, latitude b_a) on galactic latitude. For this purpose precessional corrections were first ignored. Corrections were applied for galactic rotation adopting the values given by Lindblad:[4]

$$Q = -0''.0029, \quad P = +0''.0033, \quad l_o = 327°.$$

The results for solar motion for three different galactic zones are given below:

GALACTIC ZONE	l_a	b_a	$\left(\frac{\bar{h}}{\rho}\right)$
0° to ±20°	29°.7	+11°.9	0''.0118
±21° to ±40°	33°.6	+14°.0	0''.0121
±41° to ±90°	35°.0	+15°.4	0''.0172

The probable errors in l_a and b_a are about 3°, those in $\left(\frac{\bar{h}}{\rho}\right)$ about 0''.0010. There is no indication so far of systematic differences in the position of the apex for different latitudes. Nor is any appreciable difference shown between the northern and southern galactic hemisphere:

GALACTIC LATITUDE	l_a	b_a
0° to +90°	34°.2	+15°.5
−1° to −90°	33°.9	+11°.5

Hence a combined solution was made, introducing a factor 1.43 in the coefficients of X, Y, Z in the ±41° to ±90° zones. The result is:

$$l_a = 32°.8, \quad b_a = +13°.0, \quad \left(\frac{\bar{h}}{\rho}\right) = 0''.0119.$$

With this preliminary position of the apex, values for $\left(\frac{\bar{h}}{\rho}\right)$ were computed for different galactic zones (table 1 and Fig. 1).

TABLE 1

MEAN SECULAR PARALLAX (ALL STARS) FOR DIFFERENT GALACTIC ZONES (UNIT 0''.001)

	NORTH				SOUTH			
GALACTIC LATITUDE	$\left(\frac{\bar{h}}{\rho}\right)$	P. E. ONE EQUATION	NO. OF STARS	$\bar{m}$	$\left(\frac{\bar{h}}{\rho}\right)$	P. E. ONE EQUATION	NO. OF STARS	$\bar{m}$
61° to 90°	+23.5±2.2	5.8	377	11.2	+17.6±1.6	2.9	193	11.0
41° to 60°	+15.8±1.3	4.6	1056	11.2	+14.4±1.1	3.0	913	11.4
21° to 40°	+14.7±1.4	3.9	1254	11.1	+10.0±0.9	2.4	1817	11.3
11° to 20°	+12.6±1.6	3.2	1978	11.3	+ 8.3±2.0	4.3	1733	11.5
0° to 10°	+15.1±1.6	3.3	3526	11.3	+12.4±1.2	2.9	4935	11.4

The mean secular parallax of all stars when weighted according to the number of stars in each zone is 0".0131.

To derive reduction factors the values of $\left(\frac{\bar{h}}{\rho}\right)$ just given were smoothed graphically. The ratio of this smoothed $\left(\frac{\bar{h}}{\rho}\right)$ to the mean secular parallax of all stars has been adopted as the reduction factor to be used in the solar motion terms of the general solution. Since the probable error of one equation (table 1) shows some dependence on galactic latitude, groups in the two zones +61° to +90° and +41° to +60° were assigned half weight.

Using the reduction factors and weights as described, the following values of the unknowns were derived: (Unit = 0".001)

whence $\left(\frac{\bar{h}}{\rho}\right) = 12.4 \pm 0.5$

$X = +9.99 \pm 0.50$
$Y = +6.75 \pm 0.41$
$Z = +3.00 \pm 0.45$
$Q = -3.00 \pm 0.61$
$C = +0.59 \pm 0.64$
$S = -2.93 \pm 0.60$
$\Delta' p = -0.27 \pm 0.99$
$-\Delta' e - \Delta' \lambda = +0.29 \pm 0.94$
p.e. of one equation, ±3.39

$l_a = 34°.0 \pm 2°.1$
$b_a = +14°.0 \pm 2°.1$
$Q = -3.00 \pm 0.61$
$P = +2.99 \pm 0.60$
$l_0 = 321° \pm 6°$

3. *Discussion of General Solution.*—(*a*). The position of the apex differs considerably from the one derived from the proper motions of bright stars. In equatorial coördinates the position of the apex derived here is R. A. = 19^h.0, Decl. = +36°, while the apex with respect to the bright stars is at R. A. = 18^h.0, Decl. = +30° which makes a difference of 15°. In order to have an independent check on the R. A. of the apex, two additional solutions were made for the solar motion, using proper motions in R. A. of faint stars determined in the process of parallax determinations. The corresponding values of the R. A. of the apex are 18^h.8 (Allegheny and Yale-Johannesburg) and 18^h.6 (McCormick). A higher percentage of high velocity stars[5] among the apparently faint stars may be the explanation of this well pronounced difference (Fig. 2 and section 5).

(*b*). The constants of differential galactic rotation found here are in excellent agreement with values previously found by others. A comparison is given in table 2:

TABLE 2

Q	P	l_o	AUTHOR
$-0''.0030 \pm 0''.0006$	$+0''.0030 \pm 0''.0006$	$321° \pm 6°$	van de Kamp and Vyssotsky
$-0''.0029$		$327°$	Lindblad[4]
	$+0''.0033$		Plaskett and Pearce[2]

The angular rotational velocity is: $P - Q = +0''.0060 \pm 0''.0008$. This gives a period of rotation at the sun of 220,000,000 ± 30,000,000 years. If we adopt 270 ± 42 km./sec. for the circular velocity of the sun,[6] the distance to the axis of rotation is found to be 10.5 ± 2.2 kiloparsecs. This value cannot be said to contradict seriously the value of 5.5 ± 1.2 kiloparsecs previously derived by one of us[7] from the space distribution of globular clusters when allowance for interstellar absorption is made. Neither dynamical nor photometric data at present seem to be adequate for any accurate determination of the size of our galaxy and whether or not it is exceptional when compared with other "galaxies."

(*c*). Table 3 gives a comparison of the corrections to Newcomb's precession constants found here and those of Oort[1] and Plaskett and Pearce.[2] The agreement is very satisfactory.

TABLE 3

$-\Delta e - \Delta\lambda$	Δp	AUTHOR
$-0''.0107 \pm 0''.0009$	$+0''.0101 \pm 0''.0010$	van de Kamp and Vyssotsky
$(-0''.0103 \pm 0''.0011)$	$(+0''.0098 \pm 0''.0012)$	van de Kamp and Vyssotsky (from μ_b only)
$-0''.0117 \pm 0''.0013$	$+0''.0113 \pm 0''.0013$	Oort
$-0''.0102 \pm 0''.0011$	$+0''.0094 \pm 0''.0012$	Plaskett and Pearce

A solution from the latitude equations only was made in order to see how much $\Delta' p$ and $-\Delta' e - \Delta'\lambda$ in the combined solution might have been affected by systematic cosmical errors which would presumably be larger in longitude. The results are:

$$\Delta' p = -0.64 \pm 1.22$$
$$-\Delta' e - \Delta'\lambda = +0.68 \pm 1.13$$

p.e. of one equation, 3.68.

No appreciable difference is shown between the two solutions. Moreover the larger p.e. of one equation in latitude, viz., 3.68, as compared with 3.19 in longitude, shows that as far as the present material goes the adopted theory leaves nothing more unexplained in l than in b; on the contrary it appears to account somewhat better for the facts in l than in b.

— — — —

It may be well to emphasize that the present values of Q, P, l_o, l_a, b_a, $\Delta' p$, $-\Delta' e - \Delta'\lambda$ *were all obtained from one general solution*, and their determination is therefore rather fundamental. The only assumption

involved, aside from the form of the equations, lies in the adopted system of absolute proper motions whose exclusive rôle has been that of providing some homogeneous system to which our motions were reduced; any systematic errors in this system may have affected our results. Naturally, smaller errors could have been obtained by solving for a few unknowns at a time.

The small errors of the present results are due to the almost negligible value of the cosmical errors. Even the usually bothersome large proper motions presented no problem, thanks to the large number of faint stars with small motions.

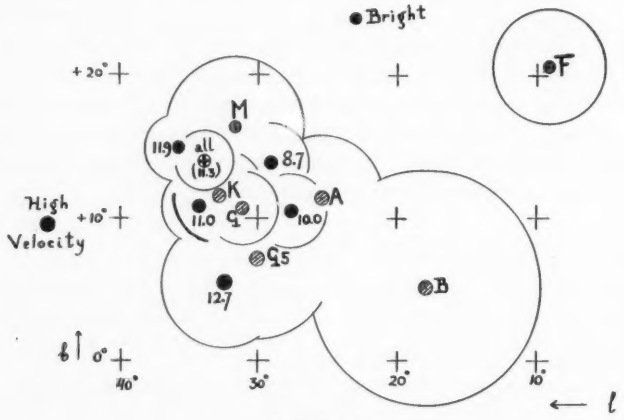


FIGURE 2

Solar apex for different magnitudes and spectra. For comparison are given the apex for stars brighter than the sixth magnitude and that for the high velocity stars.[5] The radii of the circles indicate probable errors.

It is of interest to pursue this point a little further. The probable error of one equation is 0".00339. The error of the reduction to absolute motion per region is about 0".0053. With an average of 3.24 regions per group a value of 0".00294 is found for the error in the reduction to absolute motion per group. Thus $\sqrt{(0''.00339)^2 - (0''.00294)^2} = 0''.0017$ represents roughly the probable cosmical error per group including errors of measurement, i.e., the error with which the mean motion in a group is represented by the adopted theory.

4. *Secular Parallaxes.*—As the final apex is practically identical with the provisional apex, the secular parallaxes of table 1 (Fig. 1) are considered the final values. The secular parallaxes in the northern galactic hemisphere are systematically larger than in the southern hemisphere,

the weighted mean ratio being 1.3. Moreover, the values of $\left(\frac{\bar{h}}{\rho}\right)$ in the central galactic zones are higher than in the zones immediately north or south. We believe this to be an indication of strong obscuration near the plane of the Milky Way. A further provisional investigation of the secular parallaxes for separate magnitude groups shows that in the central Milky Way zone the parallaxes of the fainter stars are considerably larger than had previously been assumed. The B, A, G and G5 stars show little dependence of secular parallax on galactic latitude. The other spectral and most magnitude groups all show a marked dependence.

5. *Solar Apex for Different Magnitudes and Spectra.*—Solutions for solar motion have been made for the different magnitude and spectral groups; the latter are of average magnitude 10. The mean motions were corrected for the galactic rotation constants derived from the general solution. No corrections for precession were applied since the general solution had given negligible corrections to the adopted precessional values.

Following the same procedure as was used for the general solution, reduction factors and weights were introduced. The computed positions of the apex are plotted in figure 2, which also gives the apex for the stars brighter than the sixth magnitude and that for the high velocity stars.

In interpreting these results it must be realized that the precessional and galactic rotation corrections were assumed, which has artificially caused a reduction in the probable errors. At the same time, however, all results depend practically on the same Boss stars, hence for differential purposes, such as studying the reality of the differences between apex positions for different star groups, the computed probable errors are maximum values.

Except probably for the F stars, there is little reason to regard the differences between the various apices as real. Conservatively, we may note a tendency toward a gradual shift in apex with increasing magnitude and also with "increasing" spectral class; the first tendency being in harmony with the suggestion already made, namely, that an increasing number of high velocity stars is present as we study fainter stars.

6. *Ellipsoidal Distribution.*—The last problem to be investigated was that of the ellipsoidal distribution of the stellar motions. Schwarzschild's method[8] was chosen because it enables one to derive constants closely connected with the constants of the theory of galactic rotation, and in addition, it permits a better use of the material as far as the regions close to the apex and vertex are concerned.

A weighted mean of two reductions from relative proper motion to absolute was adopted: (*a*) the difference between the absolute and the relative proper motion of the Boss star, (*b*) the difference between the parallactic

motion as derived here and the mean relative motion of all stars in the region. Relative weights were assigned to the two reductions according to their corresponding errors which depend on the number of Boss stars in the region and on the dispersion of peculiar motions of faint stars which is a function of galactic latitude. In addition, corrections were applied to the absolute proper motions of the Boss stars in order to allow for the new values of the precession constants and to eliminate the effect of the galactic rotation.

Several solutions for the ellipsoidal distribution were made from stars combined according to the size of their proper motions, according to galactic latitude or spectral class, and from all stars together. Table 5 gives a summary of results obtained from our material.

TABLE 5

ELLIPSOIDAL ELEMENTS FROM MCCORMICK PROPER MOTIONS

DATA	VERTEX l_v	b_v	$\alpha{:}\beta$	DATA	VERTEX l_v	b_v	$\alpha{:}\beta$
All stars	346° ± 1°6	+5° ± 1°7	1.39				
$\mu > 0''.04$	331° ± 2°8	+3° ± 2°8	....	B8–A4	341° ± 4°7	+13° ± 4°5	1.32
Δl_v	+15° ± 3°			F5–G5	340° ± 2°6	+ 6° ± 3°2	1.64
b... 0°–±20°	348° ± 2°0	+10° ± 1°9	1.41	K0–M9	343° ± 3°4	+ 2° ± 3°5	1.37
b... ±40°–±90°	337° ± 4°4	+1° ± 5°3	1.39				
Δl_v	+11° ± 5°						

It will be seen that there is a systematic difference between the longitude of the vertex l_v derived from all stars together and from stars with large proper motion; this difference has also been found in the Radcliffe proper motions[9] (our rediscussion). There is a similar difference between the vertices obtained from the stars of low and those of high galactic latitude. The change of the longitude of the vertex with the galactic latitude of the stars has also been confirmed by our rediscussion of the Radcliffe proper motions[9] and furthermore it is found in the investigations of Hufnagel[10] and Jones[11] as is shown in table 6.

TABLE 6

DATA	GAL. ZONE	l_v	GAL. ZONE	l_v	Δl_v	AUTHOR
McCormick	0°–±20°	348° ± 2°0	±40°–±90°	337° ± 4°4	+11° ± 5°	v. d. K. and V.
Radcliffe	0°–±20°	340° ± 3°4	±40°–±90°	328° ± 2°7	+12° ± 4°	Knox - Shaw and Barrett
Misc.	"galactic"	360° ± ...	"non-galactic"	346° ± ...	+14° ± ..	Hufnagel
Boss	0°–±10°	358° ± 3°0	±30°–±60°	349° ± 2°8	+9° ± 4°	Jones

Wilson and Raymond[12] have already shown that there is a systematic difference between the positions of the vertex as derived from large and small proper motions and they conclude that the position of the vertex depends on the space velocity of stars. But this effect can be equally well interpreted as a dependence of the vertex on the luminosity, or more probably on the mass of stars. This is well shown when Wilson and

Raymond group their space motions according to the absolute magnitude M of the stars:

M	l_v
≤ 3.0	$345° \pm 1°.8$
>3.0	$334° \pm 3°.9$
Δl_v	$+11° \pm 4°$

(NOTE: We have derived the probable errors from the interagreement of the subsidiary groups.)

Thus there seems to be no doubt left as to the reality of the observed effect, namely, that the vertex of the ellipsoidal distribution coincides with the direction to the center of the Galaxy for the dwarfs, but deviates from it by some 15° for the giants. It should be remarked that it is not the giants but rather the dwarfs which form the majority of the population (perhaps as much as 90% of the mass of all stars) in the vicinity of the solar system, and their motions seem to agree with the requirements of the theory of galactic rotation.

Table 5 shows that the differences between the positions of the vertex for different spectral groups are within the limits of their probable errors. The ratio of the axes of the ellipsoid, $\alpha:\beta$ is smallest for the A stars and largest for the G stars. This ratio, however, should be corrected for the accidental errors of observation. While this correction in our case is insignificant for the G stars, it would greatly increase the ratio for the A stars; unfortunately it cannot be derived for the A stars with any degree of accuracy since it is very sensitive to the adopted value of the accidental error of the observed proper motions.

For all stars together, the value of $\alpha:\beta$, approximately corrected for accidental errors, is 1.53 ± 0.03. With this value we find, using the Lindblad[13] and Oort[1] formulae, the ratio of the Keplerian force K_1 to the total force K in the vicinity of the sun:

$$K_1:K = \frac{4}{3}\left(1 - \frac{\beta^2}{\alpha^2}\right) = 0.76 \pm 0.03.$$

Plaskett and Pearce[2] find from O and B stars, $K_1:K = 0.75$, and Oort[1] 0.59 ± 0.12. It should be remarked, however, that this computation assumes the average stellar motions to be circular.

We wish to express our appreciation to Director Shapley of the Harvard College Observatory for enabling us to secure spectra for over 5000 faint stars; to Director Boss and Dr. Wilson of the Dudley Observatory for furnishing us with the improved absolute motions of 574 bright stars; to members of the staff of the Leander McCormick Observatory, especially to Dr. D. Reuyl for having measured over one-third of the proper motions, to Dr. Emma T. R. Williams for important assistance and valuable discussion, and to Director Mitchell for his continued interest and encouragement.

[1] Oort, J. H., *Bull. Astr. Neth.*, **4**, 79–89 (1927).
[2] Plaskett, J. S., and Pearce, J. A., *M. N.*, **94**, 679–713 (1934).
[3] van Rhijn, P. J., *Groningen Publ.*, No. **43**, 1–104 (1929).
[4] Lindblad, B., *Stockholm Obs. Medd.*, No. 7 (1932); *Scientia* (May, 1932).
[5] Oort, J. H., *Groningen Publ.*, No. 40, 1–75 (1926).
[6] Edmondson, F. K., *Publ. Amer. Astr. Soc.*, **8**, No. 3, 111 (1935).
[7] van de Kamp, P., *Astr. Journ.*, **42**, 97–106 (1932).
[8] Schwarzschild, K., *Göttingen Nachr., Math.-phys. Kl.*, 614–632 (1907), 191–200 (1908).
[9] Knox-Shaw, H., and Barrett, H. G. Scott, The Radcliffe Catalogue of Proper Motions (1934).
[10] Hufnagel, L., *Lund Meddelande*, **1**, No. 114, 1–31 (1928).
[11] Jones, R. D. H., *M. N.*, **91**, 563–568 (1931).
[12] Wilson, R. E., and Raymond, H., *Astr. Journ.*, **40**, 121–143 (1930).
[13] Lindblad, B., "Die Milchstrasse," *Handbuch d. Astrophysik*, **5**, 937–1076 (1933).

---

*MICROINCINERATION STUDIES. I. LOCALIZATION OF INORGANIC ELEMENTS IN PLANT CELL WALLS**

BY FRED M. UBER AND T. H. GOODSPEED

DEPARTMENT OF BOTANY, UNIVERSITY OF CALIFORNIA

Communicated June 12, 1935

Evidence regarding the structure and chemical composition of the plant cell wall has recently been presented by a number of workers. Usually the emphasis in such studies has been placed upon either ultimate physical structure or upon distribution of organic constituents such as lignin, pectin, hemi-cellulose, etc. Advances along these lines have served however, to accentuate the importance of a more accurate knowledge of the localization of mineral elements in the cell wall. For example, the "middle lamella," generally considered to be a single calcium pectate containing layer (cf. Molisch,[1] pp. 57–58), is coming to be recognized as a composite structure comprising several layers. Moreover, divergent interpretations of certain staining reactions directly involve the question of the possible presence of inorganic elements in the walls under investigation. A case in point is the ruthenium red reaction for pectin, originally advanced by Mangin[2] and now called into question by Czaja[3] who claims that the presence of calcium or phosphorus will produce a similar reaction.

It is common knowledge that the application of analytical chemical reagents to living material does not constitute a sufficient test for the presence of inorganic elements. To obviate the so-called masking which is involved, ashing or incineration of living tissue is resorted to as a standard procedure, particularly for quantitative studies. Although the

incineration of plant material and even of cut sections of woody tissues has been carried out for sundry purposes for a century, the methods employed in these early studies (reviewed in detail by Kisser[4]) were not sufficiently refined to give critical evidence on such problems as those referred to above. Only with the improved technique devised by Policard[5] and developed by Scott, Horning and others (cf. Scott[6]) for animal tissues, did the latent possibilities of the method, since known as microincineration, begin to be envisioned. It is the purpose of this communication to demonstrate the potentialities of the technique for plant anatomy and plant chemistry.

Transverse sections of the xylem of *Trochodendron aralioides*, *Betula papyrifera* and of the rotholz of *Taxodium* cut, fresh, at 5μ and preserved in 95% alcohol were very kindly furnished the authors by Dr. I. W. Bailey of the Bussey Institution. They were approximately a square centimeter in area and thus could be subdivided into smaller pieces, some of which were incinerated and others used as controls. In some cases identical sections were studied and photographed before and after incineration. The microincineration method involves (*a*) the preparation of thin sections and their subsequent mounting on standard glass slides without the addition or extraction of mineral elements, (*b*) a heat treatment which will consume the organic components, at the same time preserving the non-volatile minerals *in situ* and (*c*) satisfactory equipment for a critical examination of the residual ash.

The problem of so mounting and fixing the sections on the slides as to counteract tendencies to deformation during incineration proved difficult. At a temperature above 125°C. enormous tensions are set up in woody tissues which, unless counteracted, may cause a linear contraction of cell walls to one-fourth or less of their original dimensions. The sections referred to in what follows were held in position on the slide by a gelatin smear subsequently hardened by exposure to formaldehyde vapor. Although Eastman de-ashed gelatin was employed a detectable ash deposit usually remained and an organic adhesive entirely free from ash residue would be most desirable, especially where chemical analyses are to follow.

In order to be able to vary the ashing process over a wide range of experimental conditions, the electric oven shown in longitudinal section in figure 1 was designed and constructed. The two independent heating elements (*W*) consist of concentric helices wound on an alundum core (*A*) and, together with external resistances (*R*), and both 110 and 220 volt sources of electricity, permit a dissipation of heat within the oven varying from 400 to 4800 watts. Thus, a temperature of 600°C. can be attained either within a very few minutes or after several hours, and is indicated and controlled by a pyrometer of the potentiometer type employing chromel-alumel thermocouples (*TC*). The interior of the oven

is lined with a glazed quartz tube (*Q*) 7.5 cm. in diameter and nearly a meter long, thus furnishing a large region of constant temperature which, if necessary, can be evacuated or subjected to known atmospheric conditions. To prevent distortion of the glass slides while in a plastic state they are supported by a smooth quartz plate (*P*) during incineration.

The slides were placed in the cold furnace and the heating from room temperature to 600°C. occupied a period of 2–3 hours, the maximum temperature then being held constant for a further period of several hours, depending on the thickness, character, etc., of the material. Upon removal of the slides from the furnace, a cover slip was sealed on with a waterproof cement to protect the hygroscopic ash. Studies of the disposition and character of the ash residues, in air, were then made with the dark-field microscope preferably, although the Ultropak reflecting microscope has its special merit where the ash is exceptionally dense.

That the topography of a tissue may be described in terms of the distribution of its ash after incineration under the conditions described above

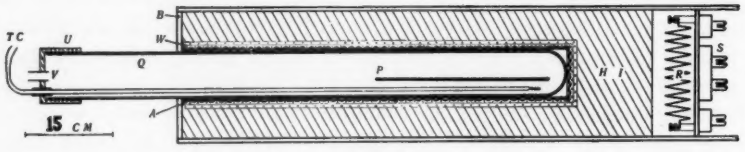


FIGURE 1

Longitudinal section of microincineration oven. *B*, transite board; *HI*, magnesia heat insulation; *W*, chromel wire, 16 B & S gauge; *U*, brass end cap. Further explanation in text.

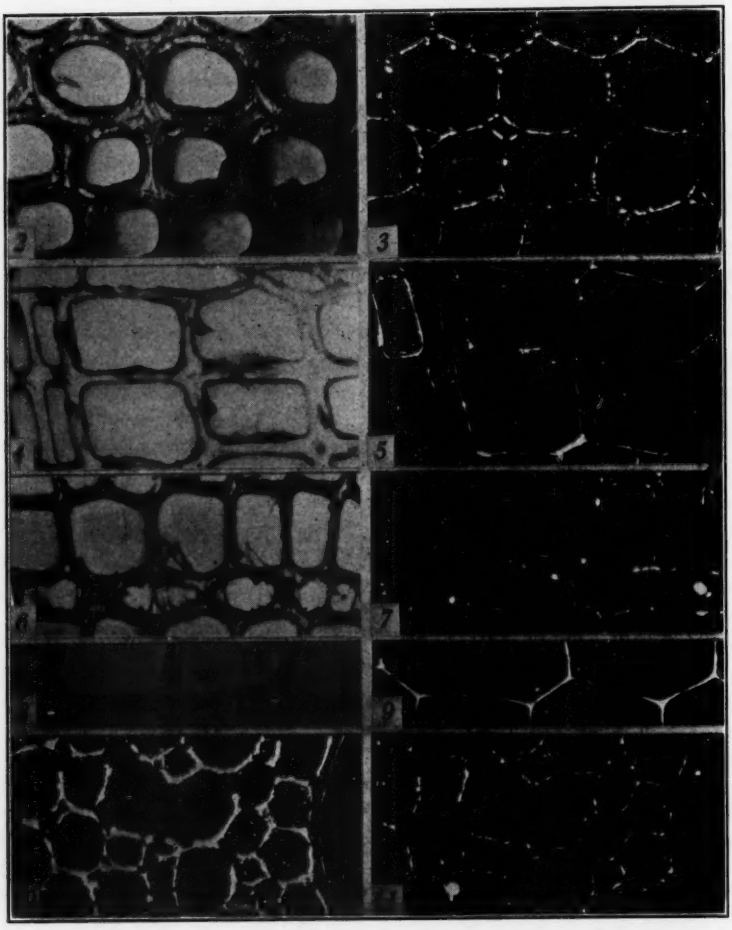
is graphically shown in the odd-numbered dark-field photomicrographs of figures 3–11. The accuracy with which the residual ash preserves the original cell patterns is made even more striking by comparing the incinerated sections with their corresponding controls in the even-numbered figures 2–10. The outstanding evidence thus revealed is the absence of conspicuous ash deposits from the greater portion of the secondary wall or, conversely, the limitation of mineral elements largely to the "middle lamella." Presumably the secondary wall may contain small amounts of non-volatile elements, but, in terms of the evidence microincineration can provide, its mineral content in the species studied is quantitatively negligible.

Adopting the nomenclature recently proposed by Kerr and Bailey,[7] the common wall between any two of the cells shown in the control photomicrographs consists of two secondary walls, two primary walls and the intercellular substance. Of these five layers at least the latter three are included in the compound "middle lamella." The question now presents itself as to whether the ash is a product of the intercellular substance, the

primary walls or the outer layer of the secondary wall. Our present evidence does not permit determination as to whether the outer layer of the secondary wall has contributed to the conspicuous ash deposits. On the question of the relative occurrence of mineral residues in intercellular substance versus primary walls, more evidence is available. A comparison of identical sections of the rotholz of *Taxodium* sp. before and after incineration (Figs. 2–3) indicates an absence of ash in the intercellular substance. Thus in figure 3 a distinct doubleness of the deposit can be seen at a number of points. A similar configuration obtains in figure 15 of Kerr and Bailey[7] which shows only primary walls following removal of the secondary walls and the intercellular material by suitable solvents. Ashed and control sections, comparable but not identical, of the summer-spring transition in *Taxodium* (Figs. 4–5) give additional evidence of doubleness in the ash deposit, but only in the late summer wood. Furthermore, the superposition of figure 3 on figure 2 indicates that the increased size of the intercellular spaces following incineration is not to be accounted for by cell shrinkage. On the assumption that the lining of the intercellular spaces (Fig. 2) consists of intercellular substance continuous with that which is common to the primary walls, figure 3 indicates that fixed minerals are not held in the intercellular spaces after incineration.

Attention is called to the ash picture of a bordered pit on the common wall between late summer and spring wood in figure 5. Assuming that no evidence of the pit membrane or torus remains and that the pit section is not median, the lenticular ash outline would represent the outer layer of the secondary wall (Kerr and Bailey, l. c., text Fig. 2). On the other hand it is possible that the ash owes its origin only to the two primary walls which compose the pit membrane, separated by ash-free intercellular substance.

Superposition of identical sections of xylem of *Trochodendron aralioides* before and after incineration (Figs. 6–7) further emphasizes the accuracy of the ash picture. An incinerated section of the same species in figure 9 reveals a coincidence of the ash deposits with the elements of the walls in the control (Fig. 8) which have taken the ruthenium red stain. It is important here to note figure 20 of Kerr and Bailey,[7] where the primary walls have been stained with ruthenium red after removal of the intercellular material with solvents for polyuronides. It will be observed that the primary walls have appreciable thickness and, in several cases, occupy the intersections to the exclusion of intercellular spaces. According to Kerr and Bailey's interpretation of the stained portions in figure 20 as primary walls, it is logical to conclude that the ash in figure 9 is a residue of the primary walls, but whether the intercellular substance and/or the outer layer of the secondary wall, contain fixed minerals also, is indeterminate.



FIGURES 2 TO 11

Control and incinerated transverse sections of xylem, 5μ. Incinerated sections photographed with bicentric dark-field condenser. Controls unstained and magnifications ×450 except as noted. Figures 2 to 5—*Taxodium* sp., (2) and (3) identical, (4) and (5) similar but not identical sections, before (2 and 4) and after (3 and 5) incineration; figures 6 to 9—*Trochodendron aralioides*, (6) and (7) identical, (8) and (9) similar but not identical sections, before (6 and 8) and after (7 and 9) incineration, (8) stained with ruthenium red, magnification (8 and 9) ×675; figures 10, 11—*Betula papyrifera*, (10) partially, (11) completely incinerated.

An incinerated section of *Betula papyrifera* is shown in figure 11 along with a photomicrograph of the identical section when partially incinerated (Fig. 10). In addition to a generally denser deposit in the latter, attention is called to a residue lining the vessel and also the fibres. It is probable that these are heavily lignified elements which burn more slowly (cf. Kerr and Bailey, l. c., Fig. 14). That they should not leave a discernible ash residue would not *a priori* be expected, since Ritter and Kurth[8] found that the ash content of lignin constituted nearly one-third of the total ash in extractive-free maple sawdust. However, it has not been demonstrated that the lignin solvents employed may not remove fixed minerals from other sources than the lignin itself.

An additional observation of interest is that the preservation *in situ* of the ash of the spring wood was achieved much more easily than in the case of summer wood. To prevent extreme distortion of the latter, especially in *Trochodendron*, an appreciably thicker smear of gelatin was required, which resulted in such heavy ash residues from the gelatin itself that interpretation was difficult. Distinction in reaction of spring versus fall wood to incineration suggests basic structural distinction between them.

The evidence presented above indicates something of the applicability of the microincineration technique to anatomical and histogenetic problems, certain of which we are further investigating by this method. Its significance in studies of the architecture and constituents of the living protoplast will be commented on in a forthcoming publication. Obviously, its ultimate importance both cytologically and histologically will depend upon the extent to which quantitative and qualitative determination of the elements present in various parts of the ash picture of a cell can be obtained. We are investigating the possibility of making such analyses.

* Investigation supported by grant from The Rockefeller Foundation.

[1] Molisch, Hans, *Mikrochemie der Pflanze.*, Jena, G. Fischer (1921).
[2] Mangin, L., *Compt. rend.*, **116**, 653 (1893).
[3] Czaja, A. Th., *Planta*, **21**, 531–604, 1934.
[4] Kisser, J., *Abderhalden's Hdb. biol. Arb.-meth.*, Abt. **11**, Teil 4, H. 2, 193–236 (1931).
[5] Policard, A., *Protoplasma*, **7**, 464 (1929).
[6] Scott, G. H., *Ibid.*, **20**, 133–151 (1933).
[7] Kerr, T., and Bailey, I. W., *Jour. Arnold Arboretum*, **15**, 327–349 (1934).
[8] Ritter, G. J., and Kurth, E. F., *Ind. & Eng. Chem.*, **25**, 1250–1253 (1933).

RELATION OF FOOD TO REGULARITY OF NUTRITIONAL RESPONSE

BY H. C. SHERMAN AND H. L. CAMPBELL

DEPARTMENT OF CHEMISTRY, COLUMBIA UNIVERSITY

Communicated May 22, 1935

Elsewhere we have described two experimental food mixtures of which the first (Diet *A*) is adequate, in the accepted sense of the word, while the second (Diet *B*) is better. In more explicit terms, Diet *A* supports normal growth and health with successful reproduction and lactation, generation after generation (rat families being still in a thriving condition after 37 generations on this diet at time of writing); but Diet *B* induces a better nutritional response in that growth is more rapid and efficient, adult vitality is superior, not only the average length of life but also the adult life expectation is improved, and the length of the period between the attainment of maturity and the onset of senility is extended in still greater ratio than is the life cycle itself.[1,2,3]

As a part of our general investigation of the nature and extent of nutritional responses to improvements of already adequate food supplies, we here compare the coefficients of variability (C. V.) observed as between strictly parallel groups of animals on these two diets, *A* and *B*. Each coefficient here cited represents data on well over one hundred animals; on several of the points of comparison the numbers were much larger.

The growth data of rats on both of these diets show so close an approximation to symmetrical frequency distribution as to justify an exceptionally high degree of confidence in the application of the usual methods of statistical interpretation.[4] Simple statistical analysis of the data of the other criteria of nutritional response briefly summarized below rests on the accepted ground that they belong to the category of natural phenomena for which, as explained by Reitz and Mitchell,[5] the assumptions of the usual statistical method may normally be regarded as valid.

1. Weight at 28 days, or growth from conception until 28 days after birth, showed a coefficient of variation (computed for all animals on each diet) of 14.3 for the rats on Diet *A*; 11.9 for those on Diet *B*.

2. Gains in weight during the 5th to 8th weeks of life (inclusive), the period of most rapid growth, were also more uniform on Diet *B* than on Diet *A*. Here the data are analyzed separately for the sexes because of the more rapid growth of the males. The growth of the males was slightly more variable than that of the females; but for each sex, growth was much more variable on Diet *A* than on Diet *B* as shown by the following figures—C. V. of males, on Diet *A*, 35.1; on Diet *B*, 21.4; C. V. of females, on Diet *A*, 33.3; on Diet *B*, 17.3.

As the animals taken for this comparison were so selected at 28 days of age that initial variability was practically the same for both diets and both sexes, the difference here shown is entirely distinct from that shown in the previous section—though both are attributable to the same difference in food supply.

3. Ages of females at birth of their first young, the sexes having grown up together, were slightly less variable on Diet *A*: C. V. for Diet *A*, 32.8; for Diet *B*, 36.2. This is probably due to the fact that some but not all of the females show strikingly earlier attainment of maturity on Diet *B*.

4. On each of the following indications of adult vitality it will be seen that gain in nutritional response on Diet *B* is accompanied also by a decreased variability or increased uniformity: (*a*) Duration of capacity for reproduction, C. V. on Diet *A*, 84.9; on Diet *B*, 65.6. (*b*) Number of young born, C. V. on Diet *A*, 87.9; on Diet *B*, 65.8. (*c*) Number of young reared, C. V. on Diet *A*, 125.1; on Diet *B*, 90.3. (*d*) Percentage of young reared, C. V. on Diet *A*, 93.4; on Diet *B*, 61.6. (*e*) Weight of young—see above.

5. The length of life was also less variable on Diet *B* though the difference here is relatively less, viz.: C. V. for males on Diet *A*, 24.4; on Diet *B*, 22.4. C. V. for females on Diet *A*, 27.7; on Diet *B*, 22.3.

Thus the main differences of response of the experimental animals to Diet *B* as compared with Diet *A*—more rapid growth, a capacity for more successful reproduction and an extension of the adult life cycle—show not only higher average attainment but greater regularity as well. Hence it would seem that in such cases as this the higher level of response is a true indication of superior nutritional well-being, and that the enrichment of the already adequate diet was here quite certainly a real improvement.

There is no necessary inconsistency between this finding and that of McCay, Crowell and Maynard[6] when due consideration is given to the fact that their experimental food mixture was very much richer in protein, and probably also in some other nutritional factors, than either of ours. When their very rich diet was fed *ad libitum* it seemed to force the rate of growth at the expense of the length of life, whereas the moderate enrichment of our initial food mixture induced *both* a more rapid and regular growth *and* an increase in the length of life. We are now seeking to establish as definitely as possible both the nature of the enrichments which constitute improvements in an already adequate diet, and the extent to which nutritional well-being is thereby improved.

The coöperation of the Carnegie Corporation of New York and the Carnegie Institution of Washington is gratefully acknowledged.

[1] H. C. Sherman and H. L. Campbell, *Jour. Biol. Chem.*, **60**, 5–15 (1924).

[2] H. C. Sherman and H. L. Campbell, *Proc. Nat. Acad. Sci.*, **14**, 852–855 (1928).
[3] H. C. Sherman and H. L. Campbell, *Jour. Nutrition*, **2**, 415–417 (1930).
[4] H. C. Sherman and H. L. Campbell, *Proc. Nat. Acad. Sci.*, **20**, 413–416 (1934).
[5] H. L. Reitz and H. H. Mitchell, *Jour. Biol. Chem.*, **8**, 297–326 (1910–11).
[6] C. M. McCay, M. F. Crowell and L. A. Maynard, *Jour. Nutrition* (in press, 1935); of which a preliminary account has been given by C. M. McCay and M. F. Crowell, *Anat. Rec.*, **57**, No. 4, Suppl. p. 102 (1933); *Nutrition Abstracts and Reviews*, **3**, 1119 (1934).

QUADRATIC WAVE EQUATION—FLOOD WAVES IN A CHANNEL WITH QUADRATIC FRICTION

By M. A. Biot

Graduate School of Engineering, Harvard University

Read before the Academy, Monday, April 22, 1935

1. *Equation of Wave Propagation.*—We shall consider a channel of constant cross-section. The depth h of the water is supposed to be the same everywhere.

If there is no friction, the displacement of a cross-sectional area of the water at time t and abscissa x satisfies the well-known wave equation (Fig. 1)

$$\frac{\partial^2 \xi}{\partial t^2} = gh \frac{\partial^2 \xi}{\partial x^2}. \tag{1}$$

The speed of propagation of the waves is $c = \sqrt{gh}$. The increase of altitude η of the water level is

$$\eta = -h \frac{\partial \xi}{\partial x}. \tag{2}$$

It satisfies the same differential equation

$$\frac{\partial^2 \eta}{\partial t^2} = gh \frac{\partial^2 \eta}{\partial x^2}. \tag{3}$$

If we assume that the channel exerts upon the moving water a frictional resistance proportional to the square of the speed $\frac{\partial \xi}{\partial t}$ a new term enters the above equation, which then becomes a quadratic hyperbolic equation.

We define a friction coefficient c_f such that the friction force per unit area is $c_f \frac{\rho}{2}\left(\frac{\partial \xi}{\partial t}\right)^2$. Then the total friction force on a cross-sectional slab

of water of thickness dx and unit width is,

$$-\frac{c_f\rho}{2}\left(\frac{\partial\xi}{\partial t}\right)^2 dx$$

where ρ is the specific mass of the water. The equation of motion is then

$$\frac{\partial^2\xi}{\partial t^2} = -g\frac{\partial\eta}{\partial x} - \frac{c_f}{2h}\left(\frac{\partial\xi}{\partial t}\right)^2.$$

By relation (2) which holds in this case too,

$$\frac{\partial^2\xi}{\partial t^2} = gh\frac{\partial^2\xi}{\partial x^2} - \frac{c_f}{2h}\left(\frac{\partial\xi}{\partial t}\right)^2.$$

Taking as variable $s = ct$ with $c = \sqrt{gh}$, this equation becomes

$$\frac{\partial^2\xi}{\partial s^2} = \frac{\partial^2\xi}{\partial x^2} - \frac{c_f}{2h}\left(\frac{\partial\xi}{\partial s}\right)^2. \tag{4}$$

This is the fundamental equation of the problem. We note that in this case this type of equation is essentially associated with the variable ξ. The altitude increment η does not satisfy a similar equation.

2. *A Solution of the Non-Linear Equation* (4).—In order to solve (4) let us try a solution of the form

$$\xi = f\left(\frac{x}{s}\right).$$

For further simplification, we prefer to put this in the form

$$\xi = \varphi\left(\frac{s+x}{s-x}\right)$$

or $\xi = \varphi(\zeta)$ with $\zeta = \dfrac{s+x}{s-x}$.

We then find

$$\frac{\partial\xi}{\partial s} = \varphi'(\zeta)\frac{1-\zeta}{s-x} \qquad \varphi'(\zeta) = \frac{d}{d\zeta}\varphi(\zeta)$$

$$\frac{\partial^2\xi}{\partial s^2} = \varphi''(\zeta)\left(\frac{1-\zeta}{s-x}\right)^2 - 2\varphi'(\zeta)\frac{1-\zeta}{(s-x)^2}$$

$$\frac{\partial^2\xi}{\partial x^2} = \varphi''(\zeta)\left(\frac{1+\zeta}{s-x}\right)^2 + 2\varphi'(\zeta)\frac{1+\zeta}{(s-x)^2}.$$

Replacing these values in the differential equation (4)

$$\zeta\varphi''(\zeta) + \varphi'(\zeta) = \frac{c_f}{8h}(1-\zeta)^2\varphi'^2(\zeta). \tag{5}$$

This is an ordinary differential equation for $\varphi(\zeta)$. It can be written

$$\frac{\varphi''}{\varphi'^2} + \frac{1}{\zeta}\frac{1}{\varphi'(\zeta)} = \frac{c_f}{8h}\left(\frac{1}{\zeta} - 2 + \zeta\right).$$

Putting

$$y = -\frac{1}{\varphi'(\zeta)}$$

$$y' = \frac{\varphi''(\zeta)}{\varphi'(\zeta)}$$

this equation becomes linear in y,

$$y' - \frac{1}{\zeta}y = \frac{c_f}{8h}\left(\frac{1}{\zeta} - 2 + \zeta\right).$$

Write $y = v\zeta$, with a new unknown v. The equation simplifies into

$$v' = \frac{c_f}{8h}\left(\frac{1}{\zeta^2} - \frac{2}{\zeta} + 1\right).$$

By integration,

$$v = \frac{c_f}{8h}\left(-\frac{1}{\zeta} - 2\log\zeta + \zeta\right) + C$$

where C is a constant of integration.

Going back through the definition of the auxiliary unknowns, we finally get

$$\varphi'(\zeta) = -\frac{1}{y} = -\frac{1}{v\zeta}$$

$$\varphi'(\zeta) = -\frac{1}{C\zeta + \frac{c_f}{8h}(\zeta^2 - 2\zeta\log\zeta - 1)}.$$

Remembering that

$$\zeta = \frac{s+x}{s-x}$$

or $$\zeta = \frac{1+\alpha}{1-\alpha} \quad \text{with} \quad \alpha = \frac{x}{s} = \frac{x}{t\sqrt{gh}}$$

$$\varphi'(\zeta) = -\frac{1}{C\dfrac{1+\alpha}{1-\alpha} + \dfrac{c_f}{8h}\left[\left(\dfrac{1+\alpha}{1-\alpha}\right)^2 - 2\left(\dfrac{1+\alpha}{1-\alpha}\right)\log\left(\dfrac{1+\alpha}{1-\alpha}\right) - 1\right]} \tag{6}$$

Putting

$$k = \frac{2}{Ch}$$

$$\varphi'(\zeta) = -\frac{kh}{2\dfrac{1+\alpha}{1-\alpha} + \dfrac{c_f k}{8}\left[\left(\dfrac{1+\alpha}{1-\alpha}\right)^2 - 2\dfrac{1+\alpha}{1-\alpha}\log\dfrac{1+\alpha}{1-\alpha} - 1\right]}$$

This value of $\varphi'(\zeta)$ is regular and perfectly defined between the values $\alpha = 0$ and $\alpha = 1$. It is not, however, the only solution of equation (5), which is also obviously satisfied by the singular solution

$$\varphi'(\zeta) = 0.$$

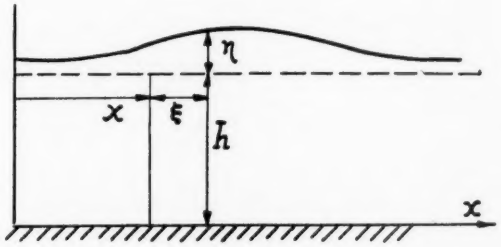


FIGURE 1

We may therefore adopt as a solution of equation (5) a discontinuous function which is

$$\varphi'(\zeta) = -\frac{kh}{2\dfrac{1+\alpha}{1-\alpha} + \dfrac{c_f k}{8}\left[\left(\dfrac{1+\alpha}{1-\alpha}\right)^2 - 2\dfrac{1+\alpha}{1-\alpha}\log\dfrac{1+\alpha}{1-\alpha} - 1\right]}$$

$$\text{for } 0 \leq \alpha \leq 1 \tag{7}$$

$$\text{and } \varphi'(\zeta) = 0 \quad \text{for } \alpha > 1$$

In order to calculate the displacement ξ of the water, we should need, of course, the function $\varphi(\zeta)$ itself. As we shall see, this is not necessary to acquire a physical picture of the solution.

3. *Physical Interpretation of the Solution.*—The elevation η of the water is given by

$$\eta = -h \frac{\partial \xi}{\partial x} = -h\varphi'(\zeta) \frac{d\zeta}{dx}.$$

By using the value (7) of $\varphi'(\zeta)$, noting that

$$\frac{\partial \zeta}{\partial x} = \frac{2}{s(1-\alpha)^2}$$

we get the relative elevation of the water between $\alpha = 0$ and $\alpha = 1$.

$$\frac{\eta}{h} = \frac{h}{s} \frac{k}{(1-\alpha^2) + \frac{c_f k}{4}\left[\alpha - {}^1/_2(1-\alpha^2)\log\frac{1+\alpha}{1-\alpha}\right]} \qquad (8)$$

For $\alpha > 1$ we get, according to (7),

$$\frac{\eta}{h} = 0.$$

This value of η represents the amplitude of a wave as shown in figure 2. For instance, for $k = 1$ and $\frac{c_f k}{4} = 0.5$, the shape of this wave will be the curve $ABCD$. At point B it has a vertical slope. The length BC is the elevation of the wave front, its relative value is given by the general formula (8) for $\alpha = 1$.

$$\frac{\eta}{h} = \frac{4}{c_f} \cdot \frac{h}{s}$$

or, since $s = t\sqrt{gh}$,

$$\frac{\eta}{h} = \frac{4}{c_f} \cdot \frac{h}{t\sqrt{gh}}.$$

It is inversely proportional to time and to the friction coefficient c_f.

The front of the wave corresponds always to $\alpha = 1$ or $\frac{x}{t\sqrt{gh}} = 1$, which means that this wave front moves with the speed

$$c = \sqrt{gh}.$$

At the origin $\alpha = 0$ or $x = 0$, the relative elevation of the wave is

$$\frac{\eta}{h} = k\frac{h}{s} = k\frac{h}{t\sqrt{gh}}. \qquad (9)$$

Note too that the area of the wave curve is independent of time. The length varies like $s = t\sqrt{gh}$ and the altitude is proportional to $\frac{1}{s} = \frac{1}{t\sqrt{gh}}$.

For instance, the shape of the wave $ABCD$ (corresponding to $k = 1$ and $\frac{c_f k}{4} = 0.5$) at various instants is illustrated by figure 3.

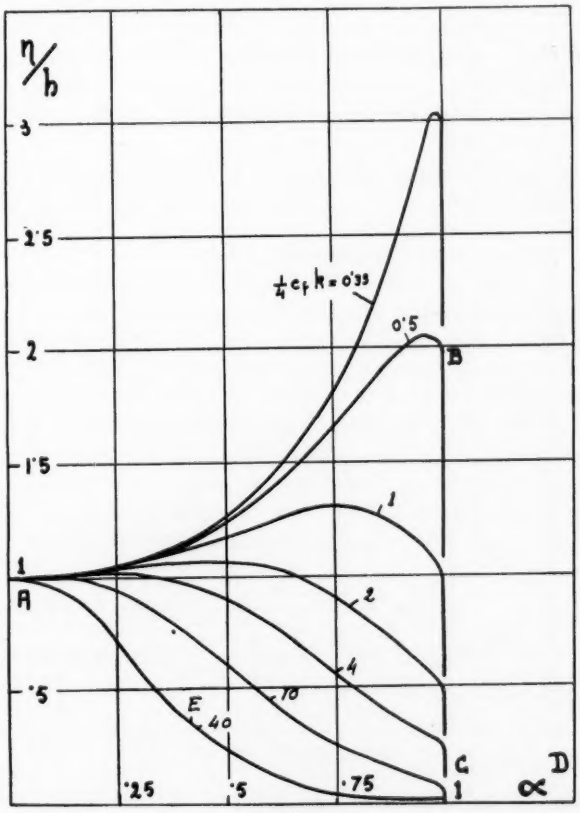


FIGURE 2

Various Shapes of Waves—$\left(\text{Shape Factor } \frac{c_f k}{4}\right)$.—In figure 2 are plotted the curves $\frac{\eta}{h}$ for a value $k = 1$ and different values of $\frac{c_f k}{4}$.

High values of $\frac{c_f k}{4}$ correspond to either high amplitudes or high friction coefficients. The effect is to damp out the wave front. This wave front still propagates with the same speed $c = \sqrt{gh}$ but its amplitude is small

compared to that of the average elevation of the wave. The energy of the wave does not propagate as fast as the wave front but only with a fraction of the wave speed. For instance in case $\frac{c_f k}{4} = 40$, point E of altitude $0.38 \frac{h}{t\sqrt{gh}}$ propagates with a speed $c' = 0.4 \sqrt{gh}$.

For low values of the friction or of the amplitude, the coefficient $\frac{c_f k}{4}$ is small. In that case, as shown by figure 2, most of the energy is kept in the wave front.

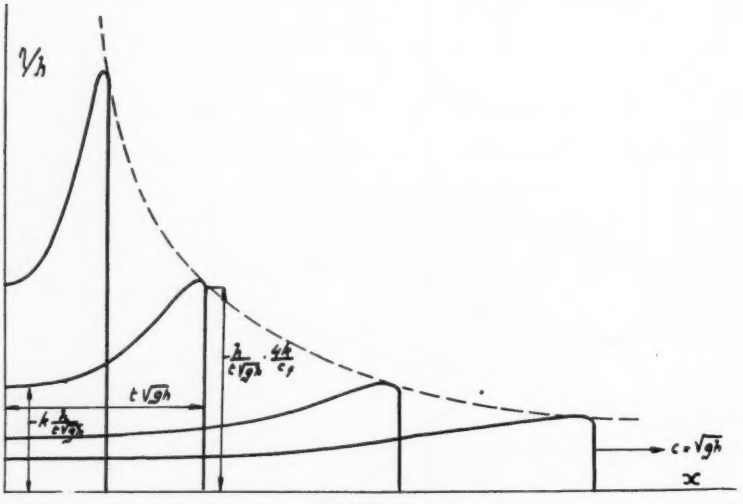


FIGURE 3

We thus get distinctly two types of waves. Waves with a steep front for values of $\frac{c_f k}{4} < 1$ and damped waves with a low front for $\frac{c_f k}{4} > 1$.

This coefficient $\frac{c_f k}{4}$, which may be called the "shape factor" determines the shape of the wave. It is noteworthy to point out the fact that this shape factor is the product of the friction by a factor k on which depends the volume or the average elevation of the wave. The physical meaning of this is clear when we remember that we have here a quadratic friction law: the higher the amplitude of the wave, the higher the speeds and hence the total friction. It is therefore in complete accordance with our natural

intuition that in cases of quadratic friction the theory shows an increase of damping effect when the elevation of the wave increases. This is a new feature in the theory of wave propagation; it is introduced by the non-linear character of our wave equarion.

The Determination of k (Volume Factor).—The coefficient c_f is supposed to be given by the physical conditions of the channel. The value of k will then depend entirely on the average elevation of the wave.

An infinite number of waves other than the ones here investigated may exist in a channel, depending on the initial elevations and speeds. It is, however, probable that if we start with a local swelling initially at rest it will propagate and deform as the wave of same average height investigated above. The origin of time to be chosen depends on the height of the wave front as shown by equation (9). The average height or volume determines the coefficient k to be used for the corresponding theoretical curve. The type of initial condition here mentioned may be roughly identified with conditions arising from a sudden input of water at a given point of a river or a channel and propagating as a flood wave.

4. *Conclusion.*—An exact solution of the equation of propagation of waves with quadratic damping has been found. It shows that high amplitude waves are more quickly damped and how this damping effect depends on both the volume of the wave and the friction coefficient of the channel. The solution may be interpreted physically as representing certain types of flood waves.

MATERNAL EFFECT AS A CAUSE OF THE DIFFERENCE BETWEEN THE RECIPROCAL CROSSES IN DROSOPHILA PSEUDOÖBSCURA

BY TH. DOBZHANSKY

W. G. KERCKHOFF LABORATORIES OF THE BIOLOGICAL SCIENCES, CALIFORNIA INSTITUTE OF TECHNOLOGY

Communicated May 27, 1935

Reciprocal crosses not infrequently produce different results; this fact is sometimes quoted as an evidence for cytoplasmatic inheritance, although the fallacy of the argument has been repeatedly pointed out. It is to be granted, however, that the cause of the difference between reciprocal crosses has not been adequately analyzed in most cases. An approximation toward such an analysis is now possible in the hybrids between race *A* and race *B* of Drosophila pseudoöbscura Frolowa.

The cross $B\ ♀ \times A\ ♂$ produces in F_1 generation sterile males having

small testes; the male offspring obtained in the F_1 of the A ♀ × B ♂ cross is likewise sterile, but the testes are normal in size.[1] The data bearing on the problem of the mechanism responsible for this difference may be summarized under the following headings.

(1). *The Difference between the Reciprocal Crosses Does Not Extend Beyond the F_1 Generation.*—Lancefield[1] found that the F_1 females are fertile, and can be back-crossed to males of either parental race. Irrespective of whether the female is an A ♀ × B ♂ or a B ♀ × A ♂ hybrid, the males resulting from the backcrosses have testes varying from normal to a very small size. The variations are due to the different combinations of the chromosomes of the parental races: a male carrying an X-chromosome of race A and a majority of the autosomes of the same race has large testes, while a male having a race A X-chromosome and race B autosomes has small testes, irrespective of whether it has race A or race B cytoplasm. The same holds true for males carrying race B X-chromosome: such males have large testes if the autosomes come also from race B, and small testes with race A autosomes. Every one of the three large autosomes present in the caryotype of the species is thus concerned with the testis size in the hybrids.[2] Experiments are now in progress to establish whether males having identical chromosomes and different cytoplasms have testes of precisely the same size, but it is already clear enough that the effect of the cytoplasm in the backcrosses, if any, is negligible as compared with the F_1 generation.

(2). *Testis Size Is Not Affected by Lack or Excess of Y-Chromosomes in Males of Pure Races.*—Schultz[3] has observed race A males having no Y-chromosome (XO). He obtained a number of such males in the offspring of normal females crossed to males carrying a translocation between the Y-chromosome and one of the autosomes. About 3% of the males in the offspring of this cross are XO (75 out of the total of 2080 males in the combined data of Drs. J. Schultz and A. H. Sturtevant). Schultz obtained also XYY males (in the progeny of XXY females). The XO and XYY conditions were checked cytologically in some cases. The XYY males have testes of normal size and are fertile. XO males are sterile, but their testes are normal in size and in histological structure. The writer obtained a single XO male in race B (in the progeny of females treated with x-ray and crossed to normal male). This male had testes of normal size; lack of the Y-chromosome was established genetically and checked cytologically.

(3). *Lack of Y-Chromosome Does Not Affect Testis Size in the Hybrids.*—Schultz[3] obtained XO hybrid males by crossing race B females to males carrying the translocation mentioned above. They had small testes, just as their XY sibs. The present writer[4] observed XO males in the backcrosses of B ♀ × A ♂ hybrid females to race A males. Their appearance is due to primary non-disjunction of the X's in the mother. These males

had testes variable in size and structure, in a manner similar to their XY sibs.

(4). *Males Coming from the B ♀ × A ♂ Cross Have Small Testes Even if They Carry the X-Chromosome of Race A.*—Orange Curly race *B* females were crossed to beaded yellow short Bare purple race *A* males. Orange and purple are autosomal recessives; Bare and Curly autosomal dominants, lethal when homozygous; beaded, yellow and short sex-linked recessives. Regular male offspring from this cross are wild-type, Bare, Curly and Bare Curly; they have small testes. Among several thousands of such males, one beaded yellow short Bare, one beaded yellow short Bare Curly and one beaded yellow short males were found. These three exceptional males are due to non-disjunction of the *X*-chromosomes in the mother, and must be XO in constitution (the XO condition was established cytologically in one of them). The testes in these males were as small as in their sibs.

(5). *Males Coming from A ♀ × B ♂ Cross Have Small Testes if They Carry the X-Chromosome of Race B.*—Race *A* females homozygous for the sex-linked recessives eosin, magenta and short were treated with x-rays, and crossed to wild-type race *B* males. Regular males in the offspring are eosin magenta short. Among about two thousand such males, one wild-type male was found. Cytologically it was found to be XO; it had distinctly small testes.

Conclusions.—The F_1 hybrid males from the *A* ♀ × *B* ♂ cross normally have the *X*-chromosome of race *A*, the *Y*-chromosome of race *B* and the cytoplasm of race *A*. The F_1 males from the cross *B* ♀ × *A* ♂ have the *X* and the cytoplasm from race *B*, and the *Y* from race *A*. In either case, half of the autosomes come from race *A* and half from race *B*. Which of these elements determine the differences between the reciprocal crosses?

Presence of a *Y*-chromosome is not necessary for the development of large testes, since large testes may be present in XO males (see paragraphs 2 and 3). An interaction between the *Y*-chromosome of race *A* and the hybrid autosomes, or race *B* cytoplasm is not necessary for formation of small testes, since small testes can be produced in other circumstances (paragraphs 1, 3, 4, 5). Small testes may be present in XO males having an *X*-chromosome of race *A* (paragraph 4) or of race *B* (paragraphs 3 and 5). The conclusion is justified that the *Y*-chromosome is not concerned with the determination of the testis size in F_1 hybrids.

Small testes are produced in F_1 hybrid males under two conditions: (*a*) if they carry hybrid autosomes in race *B* cytoplasm (paragraphs 3 and 4), it being in this case immaterial whether they carry race *A* or race *B* *X*-chromosome, and (*b*) if race *B* *X*-chromosome is present together with the race *A* cytoplasm (paragraph 5). F_1 hybrid males have large testes if race *A* *X*-chromosome is present in race *A* cytoplasm.

In short, testis size in F_1 hybrids is determined by an interaction between the chromosome complement and the cytoplasm. We are dealing either with a real cytoplasmic effect (determination by the intrinsic properties of the cytoplasm, independent of the chromosomes it carries or carried), or with a maternal effect (determination of the properties of the cytoplasm by those of the chromosomes it carried before fertilization). The data available permit a discrimination between these two possibilities. A back-cross male having one half of its autosomes from race *A* and the other half from race *B* has small testes irrespective of the source of its cytoplasm. Furthermore, back-cross males having either kind of *X*-chromosomes and either kind of cytoplasm may have large or small testes depending upon the combination of the autosomes they carry. Hence, in Drosophila pseudoöbscura, the difference between the F_1 generations of the reciprocal crosses is due to a maternal effect and not to a plasmatic inheritance.

[1] Lancefield, D. E., "A Genetic Study of Crosses of Two Races or Physiological Species of Drosophila Obscura," *Zeits. ind. Abst. Vererbungsl.*, **52** (1929).

[2] Unpublished data of the writer.

[3] Morgan, T. H., Bridges, C. B., and Schultz, J., "The Constitution of the Germinal Material in Relation to Heredity," *Carnegie Inst. Year Book*, **29** (1930).

[4] Dobzhansky, Th., "Rôle of the Autosomes in the Drosophila pseudoöbscura Hybrids," *Proc. Nat. Acad. Sci.*, **19**, 950–953 (1933).

THE DETERMINATION OF SEX IN HABROBRACON

By George D. Snell

Roscoe B. Jackson Memorial Laboratory, Bar Harbor

Communicated May 7, 1935

In the parasitic wasp, *Habrobracon juglandis* (Ashmead), as in other hymenoptera, males are normally haploid, arising from unfertilized eggs. It has been shown by Anna R. Whiting (1925), however, that following inbreeding, males may arise from fertilized eggs. These males are diploid, inheriting a chromosome set from each parent, and are referred to, therefore, as biparental (Torvik 1931). They are commonly sterile. Why some diploid zygotes should give rise to males and others to females has proved a very puzzling problem, and a problem of double interest because of its probable bearing on the normal mechanism of sex determination in the hymenoptera. In other groups of plants and animals a doubling of the whole chromosome complement does not produce changes comparable to the change of sex that it produces in the hymenoptera. Since fertilized and unfertilized eggs have the same chromosomal balance, and, in most respects at least, the same complement of genes, there is no obvious reason why the one should give rise to females and the other to males.

In considerable part, this problem has been solved by recent investigations of Dr. P. W. Whiting and his co-workers. It has been shown that femaleness is caused by heterozygosity for an allelomorphic pair of sex factors, *X* and *Y*, rather closely linked with the fused (*fu*) locus. The factors *X* and *Y* may be thought of (Whiting 1933b) as composed of several adjacent or very closely linked genes, some dominant and some recessive. Thus the *X* factor or chromosome may contain the genes *F* and *g*, the *Y* factor the genes *f* and *G*. Haploid zygotes, carrying either *X* or *Y* but not both, develop into males, and, in inbred stocks at least, the same is true of homozygous diploid zygotes, *XX* or *YY*. The lines of evidence are two.

First, when black fused males are crossed to orange females heterozygous for fused, pair matings being used, fused females and wild-type biparental males make up either much more than or much less than one-half of all the diploid progeny (Whiting 1933b and Bostian 1935). The results are explained by assuming the parent female to be of the constitution *FuX/fuY* and the parent male to be of the constitution *fuX* or *fuY*. If the parent male is *fuX*, biparental males among the progeny are *FuX/fuX* (non-crossovers) and *fuX/fuX* (crossovers). If he is *fuY*, biparental males are *fuY/fuY* (non-crossovers) and *FuY/fuY* (crossovers). There is about 15% crossing-over between the fused locus and the locus of the sex factor.

Second, the genitalia of right-left mosaic males frequently show feminization, indicating that femaleness results from the interaction between substances produced by two different haploid sets of chromosomes (Whiting, P. W., 1933a; Whiting, P. W., Greb and Speicher, 1934). Mosaic males are recognized by the presence of two types of tissue, one manifesting some given mutant character, the other failing to manifest it. They always come from mothers heterozygous for the given mutant gene, and probably are the result of the simultaneous development of a polar nucleus along with the egg nucleus. If the line of demarcation between the two types of tissue passes near the genitalia, these often show female characteristics. The assumption that one half of the body carries the *X* factor, the other the *Y* factor and that femaleness is caused by the interaction of diffusible substances produced by these two factors, will adequately explain the observed facts.

While the assumption of a single allelomorphic pair of sex factors (each perhaps a complex of several linked genes) adequately explains the above facts, it leaves one important fact unaccounted for, namely, the relative infrequency with which biparental males appear. If females are *XY* and haploid males either *X* or *Y*, half of all biparentals should be *XX* or *YY* and hence males. As a matter of fact, biparental males never make up as much as 50% of the total of all biparentals. When unrelated wasps are mated, no biparental males are produced, and even in fraternities from related wasps the females are usually at least three or four times as frequent as their biparental brothers.

To meet this difficulty, Whiting (1933b and 1935) has suggested that crosses of unrelated stocks fail to produce biparental males because of heterosyngamy. "Heterosyngamy always occurs when parents are unrelated, for all biparentals from such crosses are females. If parents are related, however, a few biparental males appear because of homoeosyngamy." In other words, *X*-bearing sperm nuclei tend to unite with *Y*-bearing egg nuclei, and vice versa, though the strength of the tendency varies with the degree of the relationship of the parents. While this suggestion will explain the low incidence of biparental males, there are a number of reasons for believing that it is not the correct explanation.

Let us first consider the incidence of biparental males in inbred stocks, and see if we find there any evidence of selective syngamy. Matings of orange #3 ♀ ♀ to related type #1 ♂ ♂ gave, in vials *a* and *b* when raised at 30°C., 233 type ♀ ♀, 21 type (biparental) ♂ ♂ and 251 orange ♂ ♂ (Whiting, P. W., and Anderson, 1932, table 4). Biparental ♂ ♂ thus made up, not 50%, but 7.7% of all biparentals. Experiments carried out by Anna R. Whiting in 1923 (see Whiting, P. W., and Anderson, 1932) showed a high mortality among eggs when the parents were from the same stock. When #3 ♀ ♀ were mated to #3 ♂ ♂, 49.1% of the eggs produced larvae and 81.8% of the larvae produced adults; when #3 ♀ ♀ were mated to unrelated #11 ♂ ♂, 74.3% of the eggs produced larvae and 77.7% of the larvae produced adults. These matings gave 118 ♂ ♂, 115 ♀ ♀; and 41 ♂ ♂, 177 ♀ ♀, respectively, a relative deficiency of biparentals (females) in the case of #3 × #3. Evidently about $43.7\% \left(\frac{74.3 \times 77.7 - 49.1 \times 81.8}{49.1 \times 81.8} \times 100 \right)$ less of the eggs from related parents are viable than of the eggs from unrelated parents. Moreover, the relative deficiency of biparentals from related parents suggests that the 43.7% is composed largely or entirely of eggs that would develop into biparentals. Let us assume that it is composed entirely of eggs destined, if they survived, to give biparental males, and calculate the expected incidence of biparentals on this basis. Using the data given above for #3 ♀ ♀ × related #1 ♂ ♂, the number of biparental males, on the above assumption, is

$$0.437(233 + 21 + 251) + 21 = 242.$$

This is approximately equal to 251, the number of females. Assuming, then, that the excess mortality in inbred stocks as compared with crossbred stocks takes place entirely among potential biparental males, we find that female zygotes and diploid male zygotes are produced with approximately equal frequency. In other words, the sex ratio is

$$XY:(XX + YY)::1:1.$$

This is the ratio expected in the absence of selective syngamy.

When other data are used, the figures do not always check as well as

they do with the above data, but this is to be expected in view of several important variable factors, particularly the highly variable incidence of haploid males. The two sets of data chosen were selected because they involve essentially the same stocks and the same environmental conditions.

The hypothesis of heterosyngamy thus appears less probable than the hypothesis of differential mortality as an explanation of the observed incidence of biparental males in inbred stocks.

It remains to determine if selective syngamy is the best explanation of the absence of biparental males in fraternities produced by unrelated parents. Certainly there is no evidence for differential mortality in this case, in fact the total mortality among eggs would be insufficient to account for the absence of biparental males if such males were initially formed with the same frequency as females.

The most probable explanation appears to be that, instead of there being a single pair of factors which, when heterozygous, causes diploid zygotes to develop into females, there are a number of such pairs, and that some of these, at least, are on different chromosomes. To take the simplest case (probably much simpler than the actual situation), assume two pairs of factors *X* and *Y*, and *W* and *Z*, lying on two different pairs of chromosomes. Females are then *XYWZ*, *XYWW*, *XYZZ*, *XXWZ* or *YYWZ*. Diploid males are *XXWW*, *XXZZ*, *YYWW* or *YYZZ*. Crosses between different stocks tend to involve different factors, for example, to be ♀*XYWW* times ♂*YZ*, and hence to give only females in the F_1. Of course if only two pairs of factors were involved, individuals from unrelated stocks would not always differ in this fashion; if five or six or more such factors are at work, however, the chance of unrelated parents giving diploid males in the F_1 becomes rather slight. Perhaps the most probable assumption is that there are a few major loci, i.e., loci which acting singly, will usually or always cause femaleness when heterozygous, and a much larger number of minor loci, i.e., loci which, acting alone, cannot cause femaleness, but which will cause it in diploid individuals where a number of the loci are heterozygous. There are two lines of evidence in support of this hypothesis.

First, it is in good accord with the known facts of sex determination in other species. It is known that in Drosophila there are multiple sex-determining factors on the *X*-chromosome (Dobzhansky and Schultz, 1934). A closer parallel is found in the fish, Lebistes. In this species, Winge (1932) has shown that there are at least two sex factors located on different chromosomes, and that in some cases one pair, in some cases the other, plays the determining rôle in sex. There is here a precedent for the assumption that the major rôle in sex determination may shift from one chromosome pair to another.

Second, and more important, the assumption of multiple sex-determining

factors in Habrobracon satisfactorily explains the appearance of biparental males following inbreeding It may be helpful in understanding what follows to note a similarity between our hypothesis and the usual hypotheses explaining hybrid vigor, particularly in their relation to inbreeding. Castle (1934) has written, "It is significant that Whiting reports the production of diploid males only when the uniting gametes are from closely related stock, a condition under which heterosis would be at a minimum." It is our assumption that the appearance of diploid males following inbreeding is due to the loss of heterozygosity for sex factors, just as the decrease of vigor following inbreeding is due to the loss of heterozygosity for vigor factors or factors in general.

Excellent experimental data on the appearance of diploid males following inbreeding have been published by Bostian (1934). Orange-eyed females were graded up to an inbred stock of b'ack-eyed males, the following system of matings being used:

P_1 ♀♀ *oo* #3 × ♂♂ *O* #11
F_1 F_1♀♀ *Oo* × their *o* sons and nephews
F_2 F_2♀♀ *oo* × ♂♂ *O* #11
F_3 F_3♀♀ *Oo* × their *o* sons and nephews
Etc.

The experiment using these stocks (the 11-*o* experiment) was carried on to generation F_{17}. In each generation, fraternities segregated rather sharply into those containing biparental males and those not containing biparental males. The percentage of fraternities containing biparental males tended to increase with each generation, the figures being, for the odd generations starting with the F_1: 38.6, 37.9, 60.7, 77.7, 91.6, 65.7, 88.8, 100. Among fraternities containing biparental males, the percentage of biparental males to all biparentals increased from 9.1 to 17.0 to 25.9 in the F_3, F_5 and F_7 generations, respectively, and then ran rather constantly near the last figure.

These results are satisfactorily explained by assuming that stock 3 and 11 differ in several non-linked sex factors similar to the *X* and *Y* factors of Whiting, females in the two stocks being, for example, *XY-WWMM* and *XXWZLL* (plus possibly several minor factors). Biparentals in the F_1 generation are all heterozygous for one or more of the factors, and hence females. With successive generations of grading up to stock #11, all the sex loci tend to become homozygous except the one for which #11 is itself heterozygous. The rate at which they tend to become homozygous is indicated by the following calculations.

Assume that the parents are

♀♀ #3 *ooXYZZ* × ♂♂ #11 *OXW* or *OYW*

where X and Y are one pair of sex factors and W and Z another pair of sex factors on another chromosome, such that diploids heterozygous for either or both pairs of sex factors are females. Assume grading up to #11 males (OXW or OYW) through successive generations by the system of matings used by Bostian. The matings are pair matings, i.e., each female mates to but a single male.

Starting with these assumptions, it may be shown that the percentage of matings which will be of types other than $XYZZ$ or $YYWZ$ times XW, or $XYZZ$ or $XXWZ$ times YW, and hence capable of producing biparental males in the odd numbered generations, will be

F_1	F_3	F_5	F_7
0	64%	80%	83%

These figures are not in quantitative agreement with Bostian's results in the early generations of the 11-*o* experiment, but they do fit the later generations, the observed figures, beginning with the F_7 generation being 60.7, 77.7, 91.6, 65.7, 88.8, 100. This may be taken to mean that more than two pairs of sex factors were present in the early generations, but that all except two of them had dropped out by about the seventh generation. In the last generation, the Z factor also probably dropped out, all females being $XYWW$, and hence capable of producing biparental males. From the nature of the segregations involved, considerable fluctuation in the proportion of females producing biparental males is expected from generation to generation, so that the exactitude of the agreement between the observed and expected results cannot be used as an argument either for or against our hypothesis. The significant thing is that the hypothesis satisfactorily explains the general trend.

Starting with the same assumptions, it may be shown that the chance that a single female in the F_3 generation will lack the Z factor is 0.33. In the F_5 generation the chance is 0.49; in the F_7 generation, 0.57. The chance increases slightly with succeeding generations. If the line was propagated by a single female in each odd numbered generation, the Z factor probably would not be carried more than two or three odd numbered generations, but it might be carried somewhat longer if several females were used in propagation. If three or four pairs of sex loci were involved, the last one, other then the X-Y locus, might well not become homozygous till the passage of a considerable number of generations. This is in accord with the results obtained by Bostian. In his 11-*o* experiment, the last sex factor introduced by the #3 stock and not present in the #11 stock appears to have been lost in the breeding females of the sixteenth generation, all breeding females of this generation having produced diploid males.

Starting with the same assumptions, and considering only fraternities

where the genetic constitution of the parents is such that biparental males can be produced, it may be shown that the expected per cent of biparentals that will be males (biparental males × 100/biparentals) is

F_1	F_3	F_5	F_7
0	39%	41%	44%

Bostian's values are much lower than these, but this is to be expected because of the high mortality among diploid males. All that can be said is that the calculated values are of the general order required to fit the facts.

The hypothesis that femaleness in Habrobracon is due to heterozygosity for one or more of a number of sex genes lying in different chromosomes is thus in accord with available data. That such heterozygosity is the only cause concerned in sex determination cannot be asserted. Quite probably the cytoplasm plays a rôle, as suggested by Castle (1934), and some influence of the environment is also probable. The evidence is strong, however, that heterozygosity for multiple sex genes is a principal agent. Further pertinent data may be obtained from either of the two following experiments.

First, linkage tests may reveal other sex factors, in addition to the *X-Y* factors, borne on other chromosomes.

Second, critical evidence can be obtained by tests of the ratio in which fused and non-fused females appear in the F_1 of the cross

$$♀♀ \textit{fu/Fu} \times \textit{unrelated} ♂♂ \textit{fu}.$$

The hypothesis of multiple sex factors predicts a 1:1 ratio of fused to non-fused females. The selective syngamy hypothesis predicts a significant departure from the 1:1 ratio.

Bostian, C. H., "Biparental Males and Biparental Ratios in Habrobracon," *Biol. Bull.*, **66**, 166–181 (1934).

Bostian, C. H., "Sex-Linkage as Shown by Biparental Males in Habrobracon," *Amer. Nat.*, **69**, 57–58 (1935).

Castle, W. E., "Possible Cytoplasmic as well as Chromosomal Control of Sex in Haploid Males," *Proc. Nat. Acad. Sci.*, **20**, 101–102 (1934).

Dobzhansky, Th., and Schultz, J., "The Distribution of Sex-Factors in the X-chromosome of *Drosophila melanogaster*," *Jour. Gen.*, **28**, 349–386 (1934).

Torvik, Magnhild M., "Genetic Evidence for Diploidism of Biparental Males in Habrobracon," *Biol. Bull.*, **61**, 139 (1931).

Whiting, Anna R., "The Inheritance of Sterility and of Other Defects Induced by Abnormal Fertilization in the Parasitic Wasp, *Habrobracon juglandis* (Ashmead)," *Genetics*, **10**, 33–58 (1925).

Whiting, P. W., "Sex-Determination in Hymenoptera," *Biol. Bull.*, **65**, 357–358 (1933a).

Whiting, P. W., "Selective Fertilization and Sex-Determination in Hymenoptera," *Science*, **78**, 537–538 (1933b).

Whiting, P. W., "Selective Fertilization," *Jour. Hered.*, **26**, 17–22 (1935).

Whiting, P. W., and Anderson, R. L., "Temperature and Other Factors Concerned in Male Biparentalism in Habrobracon," *Amer. Nat.*, **66**, 420–432 (1932).

Whiting, P. W., Greb, R. J., and Speicher, B. R., "A New Type of Sex-Intergrade," *Biol. Bull.*, **66**, 152–165 (1934).

Winge, Ö., "The Nature of Sex Chromosomes," *Proc. Sixth Int. Cong. Genetics.*, **1**, 343–355 (1932).

THE INHERITANCE AND LINKAGE RELATIONS OF KINKY COAT, A NEW MUTATION IN THE NORWAY RAT

BY ALAN L. MITCHELL

LABORATORY OF VERTEBRATE GENETICS, UNIVERSITY OF MICHIGAN

Communicated May 17, 1935

The present investigation is concerned with a study of the inheritance and linkage relations of *kinky coat*, a new recessive mutation in the Norway rat. This mutation was discovered by Dr. H. W. Feldman in a stock of Norway rats at the University of Michigan and he kindly turned it over to me for a study of its linkage relations. Since this character has not to my knowledge been described in the literature, I shall give a brief description of it. The credit for its discovery, however, belongs to Dr. Feldman.

The appearance of rats homozygous for this character is quite striking, as every hair is kinky giving the coat a distinctly curly appearance. Newborn rats that will later have a kinky coat can be distinguished from their normal sibs by their curly vibrissae.

Kinky is the second mutation of this kind to have been observed in the Norway rat, Dr. Helen Dean King (1932) having already described a dominant "curly coat" character which arose in a strain of captive gray Norway rats.

Similar mutations affecting hair structure have been found in other rodents. Castle and Nachtsheim (1933) have described three genetic types of "rex" in the rabbit, all of which are recessive and indistinguishable in appearance. As in the kinky rats, rabbits homozygous for rex can be distinguished at birth by their curly vibrissae. The coat however is plush-like rather than kinky or curly, the guard hairs being largely lacking and the other hairs shortened. The genes for two of these rex mutations lie in the same chromosome. Crew (1933) has reported a recessive "waved" coat character in the house mouse (Mus musculus). Mice homozygous for this mutation have a distinctly curly coat when about two weeks of age. Later however, the curliness disappears and it is diffieult to distinguish adult waved mice from normals.

Feldman (unpublished data) found that the new kinky coat mutation in the Norway rat behaved as a simple recessive Mendelian character. My own results, which substantiate his conclusions, are shown in table 1.

TABLE 1

F_2 RESULTS FROM CROSSES OF KINKY × NORMAL

	NORMAL	KINKY
Observed	1297	376
Expected	1254.75	418.25
Deviation	3.26	3.26
Probable error		

The F_2 ratio shown in table 1 is the summation of data from all the linkage crosses. It will be noticed that the number of kinky rats observed is slightly below that expected, the deviation being 3.26 times the probable error. A deviation of this magnitude would be expected once in 35 trials due to chance alone and may, in this case, be due to a slightly lower vitality of the kinky coat animals. These results verify those of Dr. Feldman and indicate that kinky is a simple recessive character.

TABLE 2

F_2 DATA FROM LINKAGE TESTS BETWEEN KINKY COAT (k) AND OTHER MUTANT CHARACTERS OF THE NORWAY RAT

(The expected 9:3:3:1 ratio and the deviations divided by their respective probable errors are given for each cross. X indicates the dominant allelomorph of the character tested; x its recessive allelomorph.)

CHARACTER TESTED		XK	xK	Xk	xk	TOTAL NO.
Agouti	Observed	365	100	107	29	601
	Expected	338.1	112.7	112.7	37.6	
	Dev./P. E.	3.28	1.97	0.88	2.15	
Pink eye	Observed	150	68	45	16	288
	Expected	162.0	54.0	54.0	18.0	
	Dev./P. E.	0.53	3.12	2.10	0.72	
Hooded	Observed	140	44	45	8	237
	Expected	133.3	44.4	44.4	14.8	
	Dev./P. E.	1.31	0.99	0.15	2.72	
Blue dilution	Observed	157	51	45	15	268
	Expected	150.8	50.2	50.2	16.8	
	Dev./P. E.	1.13	0.99	1.21	0.67	
Hairless	Observed	167	46	52	14	279
	Expected	157.0	52.3	52.3	17.4	
	Dev./P. E.	1.79	1.43	0.07	1.25	

The main purpose of this study was to test kinky coat for possible linkage with the following mutant characters in the rat: agouti, pink-eye (which is linked with red eye and albinism), hooded, blue dilution and

hairless. F_2 data were used and the results, including the observed ratios, expected ratios and the deviations divided by their respective probable errors are given in table 2.

The F_2 results from a repulsion cross between kinky and agouti show an excess of animals in the normal (agouti non-kinky) class, the deviation being 3.28 times the probable error. This deviation is quite possibly due to the greater viability of rats of this genetic type and may be ignored since the observed numbers in the other classes do not deviate markedly from the expected numbers. The deviation of the critical double recessive crossover class (non-agouti kinky) is only 2.15 times its probable error and it is plainly evident that the genes for agouti and kinky lie in different chromosomes.

In the linkage test with pink eye, a slight excess of F_2 animals was obtained in the pink-eyed straight-haired class but the deviation could easily be due to chance. The other observed classes fit the expected numbers very closely and linkage obviously does not exist between kinky and pink-eye.

The F_2 results from a coupling cross between kinky and hooded show a slight deficiency of numbers in the double recessive class. Since an excess rather than a deficiency of animals in this parental (kinky-hooded) class would be expected if linkage occurred between these characters, it may be concluded that kinky and hooded are not linked.

The observed F_2 ratios from the crosses of kinky with blue dilution and with hairless fit the expected 9:3:3:1 ratio very closely and in every case the deviations are less than two times their respective probable errors and not significant. Therefore the gene for kinky coat does not lie in either of the chromosomes occupied by the genes for blue dilution and hairless.

Dr. Helen Dean King, from data in press, reports that tests made by her show kinky to be independent of the curly and brown mutations which, however, are linked with each other and constitute markers for a sixth chromosome. This being so, kinky will serve as a marker for a seventh chromosome, since it has been shown by other workers that the genes for agouti, pink-eye, hooded, blue dilution and hairless are independent of each other, and I have shown that kinky is linked with none of these.

Summary.—Kinky coat, a new mutation in the Norway rat, is shown to be a simple Mendelian recessive character, and since it is not linked with agouti, pink-eye, hooded, blue dilution or hairless (nor with curly and brown—King), its gene must lie in a seventh chromosome and may be used as a marker for that chromosome.

Castle, W. E., and Wachter, W. L., "Variations of Linkage in Rats and Mice," *Genetics*, **9** (1924).

Castle, W. E., and Nachtsheim, H., "Linkage Relations of Three Genes for Rex (Short) Coat in the Rabbit;" *Proc. Nat. Acad. Sci.*, **19**, 1006–1011 (1933).

Crew, F. A. E., "Waved, an Autosomal Recessive Coat-Form Character in the Mouse," *Jour. Genetics*, **27**, (1933).

Dunn, L. C., "Linkage in Rats and Mice," *Genetics*, **5** (1920).

Ibsen, H. L., "Linkage in Rats," *Am. Nat.*, **54** (1916).

King, H. D., "Mutations in a Strain of Captive Gray Rats," *Proc. Sixth Inter. Cong. Genetics*, **2**, 50 (1932).

King, H. D., and Castle, W. E., "Linkage Studies of the Rat," *Proc. Nat. Acad. Sci.*, 390–399 (1935).

TITANOTHERE REMAINS FROM THE SESPE OF CALIFORNIA

BY CHESTER STOCK

BALCH GRADUATE SCHOOL OF THE GEOLOGICAL SCIENCES, CALIFORNIA INSTITUTE OF TECHNOLOGY

Communicated June 13, 1935

Introduction.—Until very recently, all occurrences of American titanotheres have been known from either Eocene or lower Oligocene horizons of the Rocky Mountains and western Great Plains. Absence of any trace of these mammals in Tertiary deposits located in the region to the west of the Cordilleran Province lends considerable interest to the presence of titanotheres in one of the Sespe horizons of southern California.

Remains of titanotheres at Locality 150 have been recognized for several years, since some of this material appeared early in the collection obtained during the quarrying operations at that locality. Brief mention of the occurrence of these mammals was made in preceding papers,[1] but a more extended statement was reserved in the hope that further collecting might reveal complete skull remains. Unfortunately, this hope has not been realized. The desire to refer to the group in a broad survey of the Sespe faunas and the recent discovery of titanothere material elsewhere in the far west, make it urgent at this time to attempt an identification of the type or types from the Sespe.

The rather numerous teeth and several jaws and jaw fragments, as well as foot material, of titanotheres occurring at Locality 150 contrasts markedly with the absence of such material in the collections obtained at the upper Eocene localities situated lower in the Sespe section. It substantiates a field observation that the Brontotheriidae make their first appearance or at least undergo their first noteworthy development in Sespe time during the stage represented at Locality 150.

It is a matter of some interest, particularly from the standpoint of establishing the relationship of the Sespe uppermost Eocene stage to the Duchesne River horizon of Utah, that the titanotheres occurring in the former are similar to those found in the latter. Indeed, so close is the

PLATE 1

Teleodus californicus, n. sp.

Figure 1, type specimen, left ramus and lower dentition, No. 1398, lateral view; × 3/4.

Figures 2, 2*a*, maxillary fragment with *C*, $P\frac{1}{1}$ and $P\frac{2}{2}$, No. 1834, lateral and occlusal views; × 3/4.

California Institute of Technology Collections. Sespe Uppermost Eocene, California.

resemblance between the two that there appear at present relatively few and apparently minor characters separating them specifically.

Teleodus californicus, n.sp.

Type Specimen.—Fragmentary left ramus with lower dentition, No. 1398, C.I.T. Vert. Pale. Coll., plate 1, figure 1.

Paratype.—Anterior portion of lower jaw, No. 1120.

Referred Specimens.—A third incomplete lower jaw, No. 1126, and miscellaneous lower and upper teeth, for the most part detached specimens, some of which are shown in plates 1 and 2.

Locality.—All the material referred to this titanothere comes from Locality 150 in the Brea Canyon section of the Sespe, north of the Simi Valley, California.

Characters.—$\frac{?}{2\text{–}3}, \frac{1}{1}, \frac{4}{4}, \frac{3}{3}$. Incisors with rounded, non-cingulate crowns. Third lower incisor absent in some individuals. Canines with crowns considerably reduced in size. Upper canines with crowns not noticeably enlarged at base. Very short diastema between C and $P\underline{1}$.

$P\underline{1}$ small; crown distinctly longer than wide with two external cusps, an internal cingulum, and with a cusp-like enlargement of the inner cingulum opposite forward portion of metacone (tritocone).

$P\underline{2}$ rectangular in cross-section with transverse diameter greater than anteroposterior. External surface convex on outer side of paracone and metacone. Hypocone (tetartocone) joined with protocone (deuterocone), but constriction of intervening low ridge in at least one specimen gives greater distinctness to the two cusps. Internal cingulum present or absent.

$P\underline{3}$ with outer face of paracone convex and with outer face of metacone flatter. Deep pit present adjacent to ectoloph. Hypocone (tetartocone) distinctly smaller than protocone (deuterocone) and well defined from latter cusp. Internal cingulum always absent for short distance at base of protocone; sometimes absent at base of hypocone.

$M\underline{2}$ and $M\underline{3}$ with anteroposterior diameter usually greater than trans-

---

DESCRIPTION OF PLATE 2

Teleodus californicus, n. sp.

Figure 1, $M\underline{2}$ and $M\underline{3}$, No. 1004; figure 2, $P\underline{4}$, No. 1011; figure 3, $M\underline{3}$, No. 1833; figure 4, $M\underline{3}$, No. 1095; figures 5 and 6, $M\underline{3}$, Nos. 1835, 1836; figure 7, $M\underline{1}$ and $M\underline{2}$, Nos. 1094, 1009; occlusal views; $\times$ 2/3.

Figure 8, lower canine, No. 1831, lateral view; figures 9, 9*a*, 10, 10*a*, incisors, Nos. 1838, 1837, side and inner views; figure 11, $P\underline{3}$, No. 1091, occlusal view; $\times$ 2/3. California Institute of Technology Collections. Sespe Uppermost Eocene, California.

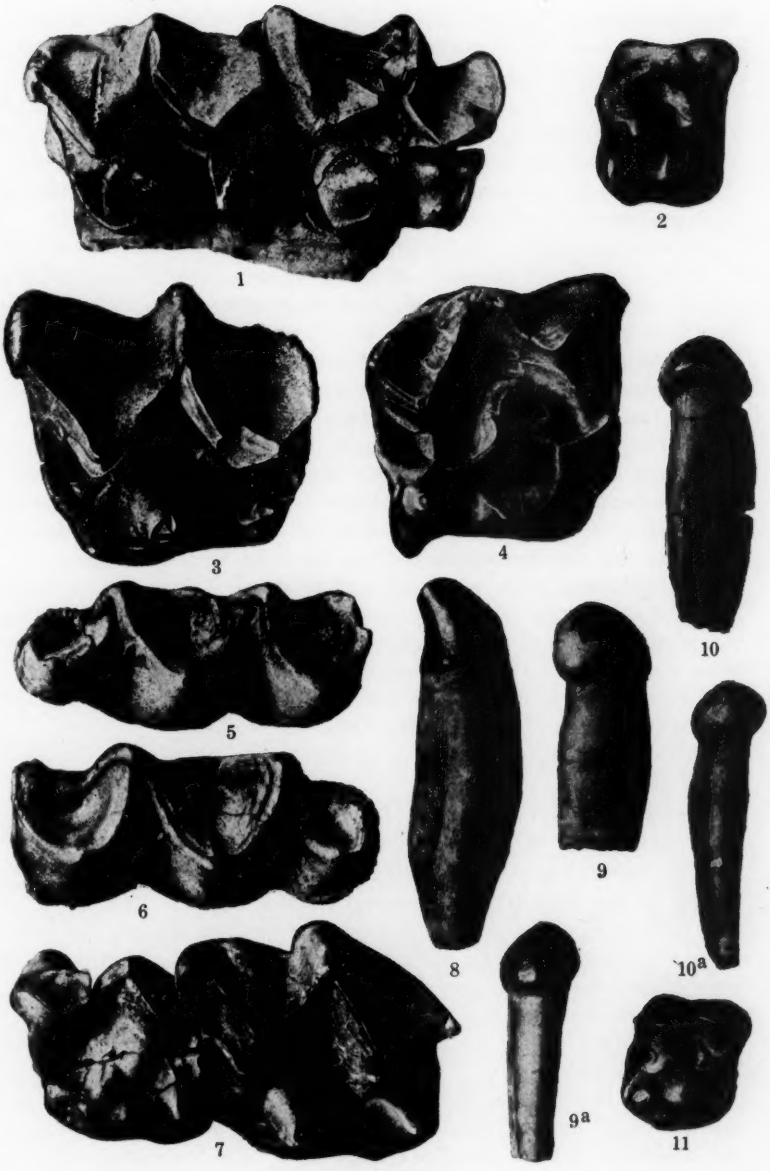


PLATE 2

(Description on opposite page.)

verse diameter. Pits form adjacent to ectoloph and hypocone is small. Hypocone in some specimens representing $M\underline{3}$ with one or more subsidiary cuspules on adjacent cingulum.

$P\overline{1}$, small.

$M\overline{3}$ may have well-developed posterior cingulum on posterior lobe.

The Sespe titanotheres are as small as or smaller than *Teleodus uintensis*.

T. californicus is distinguished principally by a precocious reduction of the lower incisors from three to two in some individuals.

Comparisons.—Reviewing the numerous types of titanotheres described from the Eocene and Oligocene of North America, the following genera, because of their geologic position, need to be especially considered with regard to possible relationship to the Sespe species, namely (1) *Protitanotherium*, (2) *Diplacodon*, (3) *Eotitanotherium*, (4) *Ateleodon* and (5) *Teleodus*. Of these, the first three are definitely Eocene in age, while the last two are from the Lower and Basal Oligocene.

(1) *Protitanotherium* differs from the Sespe form in decidedly larger size, larger size of canines, and particularly in the larger size and more primitive construction of the incisor teeth. A longer post-canine diastema is present in the skull. $P\overline{1}$ and $P\overline{2}$, relative to size of posterior premolars, are larger than the comparable teeth in the Sespe species.

(2) *Diplacodon* exhibits a number of progressive characters. In this respect it shows some resemblance to the Sespe type. Thus, for example, the premolars possess well developed tetartocones and rudimentary fossettes are present on the crown surface; the premolar tritocones are large and are subequal with the protocones, and the molars have very distinct fossettes or pits on the crown surface near the ectoloph.

Further resemblance to the Sespe form is seen in the elongate character of the molars.

However, in the Sespe species there is a shorter diastema behind the upper canine; $P\underline{2}$ has acquired a rectangular cross-section and is thus more advanced than the comparable tooth in *Diplacodon;* the premolars usually do not have complete internal cingula; and the incisors (if we may make comparison with these teeth in Peterson's species *Diplacodon superbum*) have the more advanced rounded crowns. Moreover, the lower incisors have undergone more reduction.

(3) *Eotitanotherium* is more progressive than *Diplacodon* in the characters displayed by the premolars and makes a closer approach to the Sespe form in this regard.

However, like *Protitanotherium*, *Eotitanotherium* is a large type, considerably larger than the Sespe species. Moreover, as in *Protitanotherium*, the incisors are of more primitive construction. The post-canine diastema in the skull is distinctly longer and the molars do not show the antero-posterior elongation seen in *Diplacodon* and in the Sespe form.

The three preceding types are from Eocene horizons definitely earlier in time than that represented by Locality 150.

(4) *Ateleodon*, according to Schlaikjer,[2] shows great reduction or even loss of the lower canines and incisors. In this regard the Yoder genus exhibits an advance beyond the stage represented by the Sespe specimens. In the latter the incisors and canines are present, although these teeth were on occasion small in size. The Sespe form resembles *Ateleodon* in size. It is unfortunate that the incompleteness of the material from the Yoder horizon limits severely the characterization of *Ateleodon*.

(5) *Teleodus*, or more particularly the species, *T. uintensis*, described by Peterson[3] from the Duchesne River, makes a close approach in its characters to the Sespe form.

Some specimens from the Sespe are smaller, some slightly larger than comparable materials from the Duchesne River, but on the whole there is considerable similarity in character of size between the forms from California and Utah. This difference among individuals from Locality 150 evidently expresses variation within a specific group and doubtless is due also to a difference in sex. Thus far no structural differences appear to furnish a basis for recognizing more than one specific type at the California locality.

Compared with skull No. 11754 Carnegie Mus., the maxillary fragment with C, $\underline{P1}$ and $\underline{P2}$, No. 1834 C.I.T., plate 1, figures 2, 2*a*, is similar in size with $\underline{P1}$ slightly more reduced. Two lower incisors are present in the type, No. 1398, plate 1, figure 1. A distinct difference in size prevails between these teeth, for the lateral, presumably second, incisor is considerably larger than the medial one. In the larger jaw of the paratype, No. 1120, the third incisor is likewise absent. However, in a third jaw from the Sespe, No. 1126, remnants of the forward alveolar wall and external septum suggest the presence of a small, laterally compressed alveolus for $\underline{I3}$. Unfortunately, in No. 1126, the crown of the canine immediately adjacent has been pushed into the alveolus in such a way as to damage considerably that portion of the jaw lying between the canine and $\underline{I2}$. This has prohibited a very clear demonstration of the presence of a third incisor, but there nevertheless remains, on the basis of the evidence at hand, considerable probability that $\underline{I3}$ was present. In absence or reduced size of $\underline{I3}$, the Sespe species differs from *Teleodus uintensis*. The characters displayed by the incisors in the former are certainly more like those in Oligocene titanotheres than like those in Eocene types.

Plate 2 illustrates a number of teeth representing the upper and lower dentitions of titanotheres from Locality 150. Individually, these specimens resemble comparable teeth in *T. uintensis*.

[1] Stock, C., *Proc. Nat. Acad. Sci.*, **18**, 522 (1932); *Ibid.*, **19**, 762 (1933).

[2] Schlaikjer, E. M., *Mus. Comp. Zoöl. Bull.*, **76**, 83–84 (1935).
[3] Peterson, O. A., *Ann. Carnegie Mus.*, **20**, art. 14 (1931).

DIFFERENTIABLE MANIFOLDS IN EUCLIDEAN SPACE[1]

BY HASSLER WHITNEY

DEPARTMENT OF MATHEMATICS, HARVARD UNIVERSITY

Communicated June 12, 1935

We present here a summary of some theorems on the imbedding of abstract differentiable manifolds in Euclidean space E^n and on the approximation to such manifolds by analytic manifolds. As a corollary it is noted that any differentiable manifold may be given an analytic Riemannian metric.

I. *The Imbedding of a Differentiable Manifold in Euclidean Space.*—Let M be a topological space with neighborhoods $U_1, U_2, \ldots$. Let each U_i be homeomorphic with the interior of the unit m-sphere S^m. If U_i and U_j have common points U_{ij}, then the homeomorphisms of U_i and U_j with S^m induce a mapping of one part of S^m on another part. If all such maps are of class C^r (i.e., have continuous partial derivatives through the rth order), $r \geqq 1$, with non-vanishing Jacobian, we say M is *differentiable*, and *of class* C^r.

If M is in E^n and each point of M is in a neighborhood which may be defined by expressing $n - m$ of the coördinates in terms of the remaining m, the functions being of class C^r, then M is of class C^r in the above sense; we say M *is of class* C^r *in* E^n. Suppose M is of class C^r, and is mapped into E^n. The n coördinates at points of M are n functions defined over M. If these functions are of class C^s (with the obvious definition for $s \leqq r$), and are independent (so that m independent directions at any point of M go into m independent directions in E^n), we call the map of M in E^n a *regular* C^r*-map*. Such a map is locally one-one: a neighborhood of any point of M is mapped in a one-one manner in E^n.

THEOREM I. *Any* m*-manifold of class* C^r ($r \geqq 1$ *finite or infinite*) *may be imbedded by a regular* C^r*-map in* E^{2m}, *and by such a map in a one-one manner in* E^{2m+1}.

The proof runs as follows: If M is closed, a finite number of neighborhoods $U_1, \ldots, U_\nu$ cover M. Corresponding to these neighborhoods we define functions $f_1, \ldots, f_\mu$, $\mu = (m+1)\nu$, of class C^r over M, which, used as coördinates, map M in a regular C^r-manner in E^μ. If $r = 1$, we next approximate to M by a manifold of class C^2. We now project M or the new manifold along straight lines into spaces of lower dimension, till we

have it in E^{2m+1} or in E^{2m}. If M is not closed, we define such a map successively over larger and larger parts of M.

II. *Approximations to Differentiable Manifolds by Analytic Manifolds.*—A manifold of class C^r in E^n was defined above; it is *analytic* if the functions defining its neighborhoods are analytic. If M and M^* are homeomorphic manifolds in E^n and $\eta(p)$ is a positive continuous function defined on M, we say M^* *approximates to* M *through the* r*th order with an error* $< \eta(p)$ if the distance from any point p of M to the corresponding point p^* of M^* is $< \eta(p)$, and corresponding partial derivatives of order $\leqq r$ (in a suitable coördinate system) differ by $< \eta(p)$.

THEOREM II. *Let M be of class C^r in E^n ($r \geqq 1$ finite), and let $\eta(p)$ be a positive continuous function defined on M. Then there is an analytic manifold M^* in E^n which approximates to M through the rth order with an error $< \eta(p)$.*

To prove the theorem, we first define a positive function f, analytic in $E^n - M$, and approaching 0 as we approach M. The points $f = c > 0$ define a "tube" about M; the $(n - m)$-plane orthogonal to M at p cuts $f = c$ in an $(n - m - 1)$-sphere. We define in an analytic fashion a "center" to the tube; the set of center points form M^*.

From theorems I and II follows

THEOREM III. *Any m-manifold of class C^r ($r \geqq 1$ finite) is homeomorphic with an analytic manifold in E^{2m+1}, the homeomorphism being of class C^r.*

We may define ds^2 on the manifold as the ds^2 in E^{2m+1}; hence

THEOREM IV. *Any manifold M of class C^r ($r \geqq 1$ finite) may be given an analytic Riemannian metric, the g_{ij} being of class C^r in terms of the original neighborhoods in M.*

III. *Imbedding of Manifolds in Families of Analytic Manifolds.*—We state here a generalization of Theorem II for certain classes of manifolds in E^n. We say M is in *regular position in* E^n if there exist $n - m$ independent continuous unit vector functions $v_1(p), \ldots, v_{n-m}(p)$ defined over M with the following property: Each point p_0 of M is in a neighborhood U of p_0 in M which is an m-cell, and such that any vector through two points of U makes an angle $> \rho(p_0)$ with the $(n - m)$-plane determined by the $v_i(p_0)$; $\rho(p)$ is a positive continuous function defined on M. If M is differentiable, the condition reduces to: The normal $(n - m)$-planes to points of M may be determined by $n - m$ vector functions on M. The class of such (differentiable) manifolds is the same as the class of manifolds which may be determined by the simultaneous vanishing of $n - m$ (independent) differentiable functions.[2]

THEOREM V. *Let M be an m-manifold of class C^r ($r \geqq 1$ finite) in regular position in E^n, and let $\eta(p)$ be a positive continuous function on M. Then M can be imbedded in an $(n - m)$-parameter family of manifolds M $(c_1, \ldots, c_{n-m})$, each $|c_i| < 1$, such that*

(1) $M(0, \ldots, 0) = M$.

(2) $M(c_1, \ldots, c_{n-m})$ *is analytic if* $(c_1, \ldots, c_{n-m}) \neq (0, \ldots, 0)$.

(3) *Each* $M(c_1, \ldots, c_{n-m})$ *approximates to* M *through the* r*th order with an error* $< \eta(p)$.

(4) *The manifolds fill out a neighborhood of* M *in a one-one way.*

[1] Presented to the Am. Math. Society, Sept., 1935.

[2] If M is differentiable, it is in regular position if and only if the normal sphere-space is a product space. See the following paper, especially 3 (C) and 8 (d).

SPHERE-SPACES[1]

BY HASSLER WHITNEY

DEPARTMENT OF MATHEMATICS, HARVARD UNIVERSITY

Communicated June 12, 1935

1. *Introduction.*—Spaces often occur in which the points themselves are spaces of some simple sort, for instance spheres of a given dimension. The set of all great circles on a sphere is such a space. Some general types of sphere-spaces are given in §3 below, and some specific examples in §8. Locally, sphere-spaces are product spaces (see §2); but in the large, this may no longer hold. In this note we define invariants which serve to distinguish different sphere-spaces when they have the same "base space." The proofs will be given in a later paper.

2. *Definitions.*—A sphere space is defined as follows: Let K be a topological space (usually a manifold or complex). To each point p of K let there correspond a point set $S(p)$. Let $U_1, U_2, \ldots$ be a set of neighborhoods covering K. To each U_i let there correspond a function $\xi_i(p, q)$, where p ranges over U_i and q ranges over the unit l-sphere S^l in Euclidean space E^{l+1}; for p fixed, $\xi_i(p, q)$ is a one-one map of S^l onto $S(p)$. Set $\xi_i'(p, q') = q$ whenever $\xi_i(p, q) = q'$. Suppose U_i and U_j have common points U_{ij}. Then $\xi_{ij}(p, q) = \xi_j'(p, \xi_i(p, q))$ for fixed p in U_{ij} is a one-one map of S^l onto itself; we assume that these maps are differentiable with non-vanishing Jacobian, and that they vary continuously with p. If these conditions are satisfied, we call the resulting system a *sphere-space* $S(K)$.

The space K we call the *base space.* The set of all pairs (p, q), where q is in $S(p)$, we call the *total space* $\mathfrak{S}$;[2] if the spheres $S(p)$ are non-intersecting (as parts of another space), we may let $\mathfrak{S}$ be simply the points q. $\mathfrak{S}$ is evidently a topological space.

If the maps $\xi_{ij}(p, q)$ (p fixed) are orthogonal transformations, we call the sphere-space *regular.* All the sphere-spaces described below are regular; in fact, we can define these maps by projecting the spheres $S(p)$ onto $S(p_0)$ in a simple manner for p near p_0.

3. *Some General Types of Sphere-Spaces.*—(A) Let K be the set of $(l+1)$-planes (oriented or not) through the origin in E^{n+1}. Each plane P of K intersects the unit sphere S^n in a great l-sphere $S(p)$; these spheres form a sphere-space $S(K)$.

(B) Let M be an m-dimensional Riemannian manifold. Through each point p of M there is a set of directions along M forming an $(m-1)$-sphere $S(p)$; these spheres form the *tangent sphere-space* of M.

(C) If M is a differentiable m-manifold in E^n, or in general, in a Riemannian manifold, then at each point p of M there is an $(n-m-1)$-sphere $S(p)$ of directions orthogonal to M at p; these spheres form the *normal sphere-space* of M in E^n.

(D) If M^m is a differentiable manifold in the manifold M^n, we may consider all spheres (of directions) of dimension $< m-1$ or $n-m-1$ tangent or normal to each p in M, or all such spheres at each p; the base space in this case is not M.

(E) Suppose $S(K)$ is a sphere-space. Let K' be a topological space, and ϕ, a continuous map of K' onto K. This induces a sphere-space $S(K')$: if $\phi(p')=p$, we identify $S(p')$ with $S(p)$.

4. *Orthogonal Sets of Points on a Sphere.*—We discuss here some properties of a space which are needed in the next section. Let Q_s^l be the set of all ordered sets of s orthogonal points on S^l; it is a manifold of dimension $s(2l-s+1)/2$. Let $H^t(Q_s^l)$ be the homology group of Q_s^l of dimension t. Let G^∞ be the infinite cyclic group, and G^2, the group of order 2.

LEMMA. *If $s=1$ or $l+1-s$ is even, then $H^{l+1-s}(Q_s^l)=G^\infty$; if $s>1$ and $l+1-s$ is odd, then $H^{l+1-s}(Q_s^l)=G^2$.*

The lemma holds if Q_s^l is replaced by $\overline{Q}_s^l$, defined as follows: each point of $\overline{Q}_s^l$ is a point p of S^l, together with (if $s>1$) an ordered set of $s-1$ independent directions along S^l at p.

5. *The Invariants.*—We assume here that K is an m-dimensional complex, subdivided so finely that each closed cell is in some U_i. We suppose that $S(K)$ is regular; then the orthogonality of a pair of points on any $S(p)$ has meaning. If this is not so, we would use the space $\overline{Q}_s^l$ instead of Q_s^l. Let K^r be the subcomplex of K containing all cells of K of dimension $\leqq r$. (Then $K^r=K$ for $r\geqq m$.) Let ϕ_l be a map of K° onto $\mathfrak{S}$ such that $\phi_l(p)$ is in $S(p)$. In general, let ϕ_{l-r} $(r=1,\ldots,l)$ be a map of K^r into $\mathfrak{S}$ such that $\phi_{l-r}(p)$ is in $S(p)$, and such that if p is in K^s $(s<r)$, then $\phi_{l-r}(p)$ is orthogonal on $S(p)$ to $\phi_{l-s}(p),\ldots,\phi_{l-r+1}(p)$. Such maps are easily constructed, using the fact that any singular r-sphere on an s-sphere $(r<s)$ bounds a singular cell on the s-sphere. Let ϕ be the set of maps $(\phi_0,\ldots,\phi_l)$.

Let E^r be an r-cell of K, bounded by $E^{r-1}=\Sigma E_{nj}^{r-1}$. Say E^r+E^{r-1} is in U_i. For each p in E^{r-1}, the points $\xi_i'(p,\phi_0(p)),\ldots,\xi_i'(p,\phi_{l-r+1}(p))$ form a set of $l-r+2$ orthogonal points on S^l, and thus a point $\theta_i(p,\phi)$

of Q^l_{l-r+2}. We have thus a map of the (oriented) cycle E^{r-1} onto Q^l_{l-r+2}; this determines a definite homology class of $H^{r-1}(Q^l_{l-r+2})$, and hence an element of G^∞ or of G^2. We call this element $\lambda(E^r, \phi)$, and take it as an integer or an integer mod 2. If $\Sigma\alpha_i E^r_i$ is any r-chain of K, we define

$$\lambda(\Sigma\alpha_i E^r_i, \phi) = \Sigma\alpha_i\lambda(E^r_i, \phi).$$

Suppose $L^1 = E^1_1 + \ldots + E^1_v$ is a 1-cycle of K; it is easily seen that $\lambda(L^1, \phi)$, when reduced mod 2, is independent of ϕ. Hence we may associate with each 1-cycle of K the number 0 or 1. If all such numbers are 0, we say $S(K)$ is *orientable;* otherwise, *non-orientable.* If $S(K)$ is orientable, and only then, the maps ξ_i and ϕ may be chosen so that $\lambda(E^1_i, \phi) = 0$ for each i; we shall always suppose this is done.

Suppose now that $S(K)$ is orientable; if not, we consider only chains mod 2, and we reduce all numbers mod 2. Take any r, $1 \leqq r \leqq \min(l+1, m)$, and suppose $E^r_1 \longrightarrow E^{r-1}_1 + \ldots$, $E^r_2 \longrightarrow -E^{r-1}_1 + \ldots$. Let ϕ, ϕ' be two sets of maps which differ in K^{r-1} only in E^{r-1}_1. Then it is not hard to see that

$$\lambda(E^r_1 + E^r_2, \phi) = \lambda(E^r_1 + E^r_2, \phi').$$

Hence, if L^r is any r-cycle (mod μ) ($\mu = 0$ shall mean an absolute cycle), then

$$\lambda(L^r) = \lambda(L^r, \phi) \text{ (reduced mod } \mu)$$

is independent of ϕ over $K^{r-1} - K^{r-2}$. Moreover, these numbers are independent of ϕ over K^{r-2}, and can be defined for singular cycles. Also if $L^r \approx 0$, then $\lambda(L^r) = 0$. We have therefore

THEOREM I. *Corresponding to each homology class H^r (mod μ) of K of dimension r, $1 \leqq r \leqq \min(l+1, m)$, there is a number $\lambda(H^r)$ (mod μ), which is a topological invariant of $S(K)$; if $r < l+1$ and r is even, or if $S(K)$ is non-orientable, we reduce everything mod 2.*

6. *The Invariants and the Dual Subdivision.*—Let K' represent the dual subdivision of a given subdivision of K. Say the boundary relations in K are

$$E^r_i \longrightarrow \sum_{j=1}^{\alpha_{r-1}} \eta^r_{ij} E^{r-1}_j \quad (i = 1, \ldots, \alpha_r);$$

those of the dual subdivision employ the transposed matrices $\| \eta^r_{ji} \|$. Choose a definite set of maps ϕ, and set $\gamma^r_i = \lambda(E^r_i, \phi)$. Let γ^r_i be the coefficient of the "cell"[3] of K' dual to E^r_i in a chain $L'^{(r)}$ of K'. Then by definition

$$\lambda(\Sigma\alpha_i E^r_i, \phi) = \Sigma\alpha_i\gamma^r_i = I(\Sigma\alpha_i E^r_i, L'^{(r)}),$$

the intersection of the chain in K with $L'^{(r)}$. As $L^r \sim 0$ implies $\lambda(L^r) = 0$, it follows that for each h, $\Sigma\eta^{r+1}_{hi}\gamma^r_i = 0$, that is, $L'^{(r)}$ is a "cycle" of K'.

Different choices of ϕ give different cycles $L'^{(r)}$ which are easily seen to be "homologous." Hence

THEOREM II. *For $1 \leqq r \leqq min(l+1, m)$ there is a unique homology class $H'^{(r)}$ of the dual subdivision such that if L^r is a cycle (mod μ) of K, then $\lambda(L^r) = I(L^r, L'^{(r)})$ (mod μ), $L'^{(r)}$ being a cycle of $H'^{(r)}$. If $r < l+1$ and r is even, or if $S(K)$ is non-orientable, we reduce mod 2.*

7. *Completeness of the Invariants.*—The invariants characterize regular orientable sphere-spaces in cases of low dimension:

THEOREM III. *Let K be a complex of dimension $m \leqq 3$, and let $S(K)$ and $S'(K)$ be two corresponding regular orientable sphere-spaces. If they have the same invariants, then they are equivalent in the following sense: There is a homeomorphism between the two total spaces $\mathfrak{S}$ and $\mathfrak{S}'$ such that the points of the spheres $S(p)$ and $S'(p)$ correspond. The same holds for any m if $l = 0$ or 1.*

If $m > 3$, $l > 1$, there is difficulty owing to the fact that a 3-sphere may be mapped onto a 2-sphere in such a manner that it bounds no 4-cell on the 2-sphere.[4]

COROLLARY. *If in the above theorem ($m \leqq 3$ or $l \leqq 1$) all invariants vanish, then $\mathfrak{S}$ is the product of K and S^1.*

The converse is obviously true.

8. *Examples.*—(*a*) Let $\mathfrak{S}'$ be the product of the segment (0, 1) and the circle given by the angle θ. If we identify the two ends by setting $0 \times \theta = 1 \times (-\theta)$, we obtain a 1-sided ring surface $\mathfrak{S}$. This is the total space of the 1-sphere-space $S(K)$ with K = a 1-sphere (= (0, 1) with ends identified) and with $\lambda(K) = 1$.

(*b*) The tangent sphere-space to an m-manifold M is an $(m-1)$-sphere-space with base space M. The invariant $\lambda(M)$ of highest dimension m is the same as the index sum of singularities of a vector distribution on M; as is well known, this equals the Euler-Poincaré characteristic of M. Thus if M is the 2-sphere, the tangent space (which is the same as the space mentioned in the introduction) has the invariant $\lambda(M) = \pm 2$. The total space is the same as the group of rotations, or, projective 3-space.

The lower-dimensional invariants of these spaces do not seem to have been studied.

(*c*) The tangent sphere-space of the projective plane is non-orientable. It has non-zero invariants (mod 2) of dimensions 1 and 2. The defining cycles of the dual are: $L'^{(1)}$ = a non-bounding closed curve, $L'^{(2)}$ = a single vertex.

(*d*) The normal sphere-space to any m-cell[5] or curve or orientable 2- or $(n-1)$- or closed $(n-2)$-manifold in E^n is a product space. The normal sphere-space to any manifold in E^n is orientable if and only if M is orientable.[6]

(*e*) There is an open 3-manifold M^3 in E^5 containing a 2-sphere S^2 such

that for the normal sphere-space $S(M)$, $\lambda(S^2) = 2$. If $S'(M)$ is the normal sphere-space of M in E^n, $n > 5$, then $\lambda'(S^2) = 0$. For the tangent sphere-space, $\lambda(S^2) = 0$.

(*f*) There is an open M^4 in E^6 containing an S^2 such that for the normal space, $\lambda(S^2) = 1$; this holds also if we imbed E^6 in E^n ($n > 6$). For the tangent space, $\lambda(S^2) = 1$. By the methods in the preceding note, we may extend M^4 to a closed differentiable manifold in E^9.

9. *Normal Sphere-Spaces and Intersections.*—Let M^m be a manifold of class C^2 in E^n, and let L^r be an r-cycle of M^m composed of differentiable cells of M^m, $r \leqq n - m$. By considerations of dimensionality we see that E^n may be projected into an E^{m+r} so that a neighborhood N of L^r in M^m goes into E^{m+r} in a locally one-one way, and in a one-one way if $r < m$ (see the preceding note). There is now a well defined intersection $I(M^m, L^r)$ in E^{m+r}, or a "local" intersection I_{local} (in general different from the ordinary intersection) if the mapping is only locally one-one in E^{m+r}.

THEOREM IV. *If $S(M)$ is the normal sphere-space of M^m in E^n, then $\lambda(L^r) = I$ or $I_{local}(M^m, L^r)$ in E^{m+r}; if $r < n - m$ and r is even, or if $S(M)$ is non-orientable, we reduce mod 2.*

10. *Plane Spaces.*—Suppose that to each point p of K corresponds an $(l + 1)$-plane $P(p)$, forming a *plane-space* $P(K)$, using a definition much as in §2. It is easy to pick out a point $\psi(p)$ on each $P(p)$ in a continuous manner over K. Through each $\psi(p)$ there is an l-sphere $S(p)$ of directions; we form thus a sphere-space $S(K)$. $P(K)$ may be studied by studying $S(K)$.

[1] Presented to the Am. Math. Society, April, 1935.

[2] The total spaces of some sphere-spaces have been studied by H. Hotelling, *Trans. Am. Math. Soc.*, **27**, 329–344 (1925) and **28**, 479–490 (1936), and by W. Threlfall, *Jahresb. Deutch. Math. Ver.*, **42**, 87–110 (1932).

[3] If K is not a manifold, these will not be true cells, and they may be of various dimensions.

[4] H. Hopf, *Math. Annalen*, **104**, 637–665 (1931).

[5] See W. Wazewski, *Compositio Math.*, **2**, 63–68 (1935).

[6] Compare Seifert and Threlfall, *Lehrbuch der Topologie*, pp. 272–276.

GROUPS WHICH ARE THE PRODUCTS OF TWO PERMUTABLE PROPER SUBGROUPS

BY G. A. MILLER

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ILLINOIS

Communicated June 13, 1935

Let G represent a non-regular transitive substitution group of degree n and let G_1 represent one of its subgroups composed of all its substitutions which omit a given letter. If G involves a transitive proper subgroup H then H has a subgroup of index n in common with G_1 and hence $G = G_1H = HG_1$. That is, *a necessary and sufficient condition that a transitive substitution group is the product of two permutable proper subgroups such that one of them is composed of all its substitutions which omit a given letter is that the group contains a proper transitive subgroup.* If the latter subgroup is regular then these two permutable subgroups have only the identity in common and vice versa. In particular, if n is a power of a prime number then H can always be so selected that it is regular since a Sylow subgroup of G whose order is a power of this prime number must then contain a regular group, as results from the following considerations.

Such a Sylow subgroup is transitive since the subgroup composed of all its substitutions which omit a given letter is of index n because the order of this Sylow subgroup is n times the order of the corresponding Sylow subgroup in G_1. It must therefore contain an invariant subgroup of degree n and of prime order p. The n/p transitive constituents of this subgroup represent systems of imprimitivity of this Sylow subgroup and hence this Sylow subgroup is isomorphic with a transitive group of degree n/p which also has an invariant subgroup of order p. If this is not a $(p, 1)$ isomorphism then to a substitution of order p in the latter invariant subgroup there correspond in the given Sylow subgroup substitutions which are commutative with only p of the substitutions in the subgroup corresponding to the identity of this quotient group and involving no more than p^2 letters. Hence G contains at least $p^2 - 1$ regular substitutions of degree n such that their p^{th} powers appear in the first of the given invariant subgroups of order p. As this process may be repeated when $n > p^2$ it has been proved that *every transitive group of degree p^m, p being a prime number, contains a regular group of order p^m and is therefore the product of two permutable proper subgroups which have only the identity in common, whenever it is non-cyclic.*

Every solvable group which is not both cyclic and also of prime power order is the product of two of its permutable proper subgroups since such a group contains an invariant subgroup of prime index and hence this subgroup and a Sylow subgroup whose order is a power of the same prime

number are two permutable subgroups which satisfy the required conditions whenever the order of the group is not a power of a prime number. When this order is such a power and the group is non-cyclic it was proved above that the group is the product of two permutable proper subgroups which have only the identity in common. If G contains a subgroup which is not transformed into itself by any operator of G_1 besides the identity and if this subgroup has n conjugates under G then G is the product of two permutable proper subgroups which have only the identity in common since it involves a regular subgroup of degree n. In particular, the simple group of order 660 can be represented as a transitive substitution group of degree 12 which satisfies the given conditions with respect to a cyclic subgroup of order 6.

If there is a group which satisfies the conditions that it cannot be represented as the product of two permutable proper subgroups and is not both cyclic and of prime power order there must be a simple group of composite order which also satisfies these conditions. This simple group could not belong to the system composed of the alternating substitution groups since such a group is obviously the product of two permutable proper subgroups having only the identity in common whenever its degree is not the double of an odd number, and when its degree is of this form it is the product of two permutable subgroups, one of which is composed of all its substitutions which omit a given letter and the other of the imprimitive group obtained by extending the positive substitutions in the direct product of the two symmetric groups of degree $n/2$ by a positive substitution of order 2 which permutes its systems of intransitivity.

If a group is the product of two permutable subgroups which have only the identity in common then the sum of the degrees of transitivity of these two factor groups is equal to the degree of transitivity of the original group. By means of this theorem it is easy to prove that the alternating group of degree 6 is not the product of two permutable subgroups which have only the identity in common since one and only one of these two subgroups would involve operators of order 5. This subgroup could not be simply transitive since the other factor subgroup could not be as much as three times transitive. On the contrary, the alternating group of degree 10 is the product of the alternating group of degree 7 and the three times transitive group of degree 10 and of order 720 which is a subgroup of the five times transitive group of degree 12 discovered by E. Mathieu. From this five times transitive group it results directly that the alternating group of each of the five degrees 8 to 12 is the product of two permutable subgroups which have only the identity in common and that one of these subgroups in each of these cases can be so selected that it is the alternating group of degree 7.

It was noted above that if an alternating group is not the product of

two permutable subgroups which have only the identity in common then its degree is the double of an odd number, but that this condition is not sufficient since the alternating group of degree 10 is the product of two such permutable subgroups. If an alternating group is the product of two permutable subgroups which have only the identity in common then one and only one of these subgroups must be an alternating group of lower degree since a transitive group of degree n which does not include the alternating group of this degree cannot be more than $n/3 + 1$ times transitive. Moreover, when n is of the form $2(2n + 1)$ the alternating group of degree n is not the product of two permutable subgroups which have only the identity in common if one of these subgroups is the alternating group either of degree $n-1$ or of degree $n-2$, since the other factor subgroup could not involve the regular group of degree n because this involves negative substitutions.

From what precedes it results that two fundamental questions relating to the subject under consideration are whether a given group G is the product of two permutable proper subgroups and whether these subgroups can be so selected that they have only the identity in common. No group besides the cyclic group of prime power order seems to be known which does not satisfy the condition that it is the product of two permutable proper subgroups but many such groups are not the product of two such subgroups having only the identity in common. The quaternion group is one of the simplest among those which belong to the latter category. In view of the fact that the five times transitive group of degree 12 is a well-known simple group it may be noted here that it is the product of two permutable proper subgroups since it contains a transitive proper subgroup composed of its substitutions which transform a certain substitution of order 2 and of degree 12 into itself as results from the following equations:[1]

$$s_1 = abdihfjcekl,\ s_2 = abcd.efgh,\ s_1s_2 = acfjdieklb.gh,$$
$$(s_1s_2)^5 = ai.bd.ce.fk.gh.jl,\ (s_1s_2)^5s_2^2 = aicyfkhe.jl.$$

The fourth power of the last of these substitutions and s_1s_2 generate a transitive group under which $(s_1s_2)^5$ is invariant. This cannot be the entire group since the latter is simple. In closing it may be desirable to note that the term "permutable groups" is incorrectly defined under the entry "permutable" in the second edition of Webster's *New International Dictionary*, 1935, since every finite group is unchanged when it is operated upon by the group. The illustration which follows this definition and relates to an operator which is permutable to a given group is, however, correct, but it does not apply to the preceding definition because two groups may be permutable even when neither of them transforms the other into itself. See also the entry "distribution" in this dictionary.

The terms permutable and commutative are used interchangeably when applied to operators but not always when applied to subgroups as results from the entry "commutative groups" in the dictionary to which we have just referred. There seems to exist no good reason for this difference in usage.

[1] G. A. Miller, *Bull. Amer. Math. Soc.*, **1**, 255–258 (1895).

ABSTRACT ABSOLUTE VALUES WHICH GIVE NEW IRREDUCIBILITY CRITERIA

BY SAUNDERS MAC LANE

DEPARTMENT OF MATHEMATICS, HARVARD UNIVERSITY

1. *Introduction.*—The Schöneman-Eisenstein theorem has in the past given rise to a number of more general irreducibility criteria. O. Ore[1] and J. Kürschak[2] have obtained such criteria by using Newton polygons. T. Rella[3] has pointed out that the polygons depend essentially upon the use of a "non-Archimedean" absolute value for the ring of all polynomials with rational coefficients. This suggests that other absolute values in this ring would give rise to new irreducibility conditions for polynomials. The present paper is concerned with the determination of all such values. We construct all possible non-Archimedean values in any ring $K[x]$ of polynomials in x with coefficients in a field K, provided K is assumed to have only "discrete" values. These absolute values give rise to a sequence of irreducibility theorems, containing and extending the known criteria of the Newton polygon type.

2. *The Construction of Absolute Values.*—A ring S has a non-Archimedean value V if to every element $b \neq 0$ of S there is assigned a unique real number Vb, with the properties

$$V(a + b) \geqq \text{Min } (Va, Vb) \tag{1}$$

$$Vab = Va + .Vb. \tag{2}$$

We can assume $V(0)$ to be $+\infty$. It is convenient to say that a is *equivalent* to $b (a \sim b)$ if

$$V(a - b) > Va = Vb,$$

and that a is *equivalence divisible* by b if there is a c such that

$$a \sim bc.$$

If S is a field, the set of all real numbers of the form Va is an additive group, called the *value group*.

If any such value V for a field K is given, we can construct a corresponding value V_0 for the ring $K[x]$ by considering x as an "invariant element" in the sense of A. Ostrowski.[4] This means that x is to be assigned some value μ, while any other polynomial is to have a value

$$V_0(a_n x^n + a_{n-1}x^{n-1} + \ldots + a_0) = \operatorname{Min}(Va_n + n\mu,\ Va_{n-1} + (n-1)\mu,\ \ldots,\ Va_0).$$

The axiom (1) indicates that the quantity on the right could be larger; in fact, other values can be obtained by increasing the value of some "key" polynomial $\varphi(x)$. Whenever a value V for $K[x]$ is given, a new value V' may be obtained from a key polynomial satisfying the conditions:

I $V'\varphi(x) = \mu' > V\varphi(x)$;

II If a product is equivalence divisible (in V) by $\varphi(x)$, then one of the factors is so divisible by $\varphi(x)$;

III If a polynomial $a(x)$ is equivalence divisible by $\varphi(x)$, then the degree of $\varphi(x)$ does not exceed that of $a(x)$.

Any polynomial can then be expanded in the form

$$f(x) = \sum_{i=0}^{n} f_i(x)[\varphi(x)]^i, \tag{3}$$

where the polynomials $f_i(x)$ are of smaller degree than $\varphi(x)$. The value V' is defined by

$$V'f(x) = \operatorname{Min}[Vf_i(x) + i\mu']. \tag{4}$$

Following (2), we can then define the value of a rational function in $K(x)$ thus:

$$V'(f(x)/g(x)) = V'f(x) - V'g(x). \tag{5}$$

THEOREM 1. *If V is an absolute value, and if $\varphi(x)$ and μ' satisfy conditions I to III above, then the function V' defined by* (4) *and* (5) *is an absolute value of $K[x]$ and of the field of rational functions $K(x)$.*

If we start with the V_0 first defined and apply the process repeatedly with a succession of key polynomials, $\varphi_1(x)$, $\varphi_2(x)$, ..., we obtain a value which may be symbolized thus

$$V_k = \begin{pmatrix} V & x & \varphi_1(x) & \ldots & \varphi_k(x) \\ & \mu & \mu_1 & \ldots & \mu_k \end{pmatrix} \tag{6}$$

This we call an *inductive* value. In the important arithmetic applications all the coefficient values V have discrete value groups and hence are said to be *discrete*. In this case there is a completeness theorem:

THEOREM 2. *Let W be any non-Archimedean value of $K[x]$, and assume*

all values of K to be discrete. Then either W is an inductive value of the form (6), *or else there is a sequence* V_k *of such values, so that for any polynomial* $f(x)$

$$Wf(x) = \lim_{k \to \infty} V_k f(x).$$

Here the inductive values V_k *on the right are all formed from a fixed V, a fixed* μ *and fixed infinite sequences* $\varphi_k(x)$ *and* μ_k.

3. *The Structure of Inductive Values.*—In any field with a value, the elements of positive value form a prime ideal P in the ring B of elements with non-negative values. The resulting residue-class field B/P is known to be of central importance. The structure of these fields for the values defined above may be described thus:

THEOREM 3. *If a field* $K(x)$ *has a discrete inductive value V and a residue-class field* $F(y)$ *for this value, then to any polynomial in* $K[x]$ *there corresponds a certain polynomial in* $F[y]$. *If* $\varphi(x)$ *is a key polynomial satisfying conditions I to III, then the corresponding polynomial* $\psi(y)$ *is irreducible in* $F[y]$, *and a field* F' *may be formed by adjoining a root of* $\psi(y)$ *to F. Then the residue class field for the new value* V' *defined by* (4) *and* (5) *is either* $F'(y)$ *or simply* F', *according as there is or is not an integral multiple of* $\mu' = V'\varphi(x)$ *in the value group of V.*

The proof of this theorem centers about the properties of the correspondence between $K[x]$ and $F[y]$.

4. *Irreducibility Conditions.*—Corresponding to the expansion (3) over a value V we can construct a Newton Polygon and a corresponding irreducibility theorem. This can be done for any inductively defined value V_k; the known theorems are essentially those for $k = 1$.

THEOREM 4. *Let* V' *be a value defined as in Theorem* 1. *If a polynomial* $f(x)$ *has an expansion* (3) *in which*

1. $V'f_n(x)[\varphi(x)]^n = V'f_0(x) \leqq V'f_i(x)[\varphi(x)]^i$,
2. $f_n(x)$ *is a constant,*
3. n *is the smallest integer such that* $n\mu'$ *belongs to the value group of V;*

then $f(x)$ *is irreducible in* $K[x]$.

The proof of these theorems and of some of their extensions will be published shortly in a mathematical journal.

[1] O. Ore, "Zur Theorie der Irreduzibilitätscriterien," *Math. Zeit.*, **18**, 278–288 (1923).

[2] J. Kürschak, "Irreduzible Formen," *Journ. f. d. Math.*, **152**, 180–191 (1923).

[3] T. Rella, "Ordnungsbestimmungen in Integritätsbereichen und Newtonsche Polygone," *Ibid.*, **158**, 33–48 (1927).

[4] A. Ostrowski, "Untersuchungen zur arithmetischen Theorie der Körper," Teil III, *Math. Zeit.*, **39**, 321 (1934).

ON A METRIC PROPERTY OF FUCHSIAN GROUPS

BY W. SEIDEL

DEPARTMENT OF MATHEMATICS, HARVARD UNIVERSITY

Communicated June 14, 1935

1. A Fuchsian group Γ is a properly discontinuous group of linear transformations

$$z' = \frac{\alpha z + \beta}{\gamma z + \delta}, \quad \alpha\delta - \beta\gamma \neq 0, \tag{1}$$

of the complex z-plane, each of which carries a certain circle into itself and carries each part into which the circle divides the plane into itself.[1] The invariant circle is called the principal circle and will be assumed throughout this paper to be the unit circle $|z| = 1$. It will be denoted by E.

Consider now any set of points S on E. In general, such a set will be carried by the transformations of a given group Γ into other sets. Let us suppose that a group Γ possesses the property that every measurable set of points S on E which is invariant under all transformations of Γ is either of Lebesgue measure 0 or 2π. Every such group will be called *metrically transitive on the principal circle.*[2] Any function of a complex variable $f(z)$ defined in $|z| < 1$ for which

$$f\left(\frac{\alpha z + \beta}{\gamma z + \delta}\right) = f(z)$$

for every transformation (1) of a Fuchsian group Γ will be called *automorphic with respect to* Γ.

2. THEOREM. *A necessary and sufficient condition for a Fuchsian group* Γ *to be metrically transitive on the principal circle* E *is that every bounded harmonic function* $u(z)$ *defined in* $|z| < 1$, *automorphic with respect to* Γ, *shall be identically a constant.*

The condition is necessary. It suffices to assume that there exists a bounded harmonic function $u(z)$ in $|z| < 1$ which is automorphic with respect to Γ and prove that $u(z)$ is identically a constant. Suppose that $u(z)$ is such a function and denote the transformations of the group Γ by $T_1(P)$, $T_2(P)$, ..., $T_n(P)$, ..., P being any point of the closed circle $|z| \leqq 1$. Since $u(z)$ is bounded and harmonic in $|z| < 1$, it follows from a well-known theorem of Fatou[3] that

$$\lim u(z) = f(e^{i\theta}), \quad z = re^{i\theta}, \quad 0 \leqq r < 1, \quad z \longrightarrow e^{i\theta}$$

for all points $e^{i\theta}$ of the circumference E, except perhaps for a set of measure 0, the limit being taken along paths non-tangential to E. Thus, the

function $f(e^{i\theta})$ is defined for almost all points of E and is bounded and measurable. Since $u(z)$ is automorphic with respect to Γ, $f(e^{i\theta})$ is likewise invariant under all transformations of Γ. The function $u(z)$ may now be represented by the Poisson integral

$$u(z) = \frac{1}{2\pi}\int_0^{2\pi} f(e^{it})\,\frac{1-r^2}{1+r^2-2r\cos(\theta-t)}\,dt,\ z = re^{i\theta}. \qquad (2)$$

It will be proved that if Γ is metrically transitive on E, $u(z)$ must be identically equal to a constant. Indeed, consider the set $\Omega_{a,b}$ on E which consists of all points $e^{i\theta}$ for which $a < f(e^{i\theta}) \leqq b$ for any pair of real numbers a and b such that $a < b$. The set $\Omega_{a,b}$ is measurable. Denoting by $T_n(\Omega_{a,b})$ the set on E into which $\Omega_{a,b}$ is transformed by the transformation $T_n(P)$ of the group Γ, it is seen at once that $T_n(\Omega_{a,b}) = \Omega_{a,b}$. This shows that $\Omega_{a,b}$ is a measurable set of points invariant under all transformations of the group Γ. Since Γ is metrically transitive on E, the measure of $\Omega_{a,b}$ is either 0 or 2π. It follows easily from this that there exists a constant ω such that $f(e^{i\theta}) = \omega$ for almost all points of the interval $0 \leqq \theta \leqq 2\pi$. Hence by (2), $u(z) = \omega$ identically in the whole unit circle $|z| < 1$. The necessity is, therefore, proved.

The condition is sufficient. Suppose, in fact, that the group Γ is not metrically transitive on E. Then, it suffices to construct a bounded harmonic function $u(z)$, not identically a constant, in the circle $|z| < 1$ which is automorphic with respect to the group Γ. If Γ is not metrically transitive on E, there must exist a measurable set Ω which is invariant under all transformations of Γ and such that $0 < m\,\Omega < 2\pi$. Let $f(e^{i\theta})$ be the characteristic function of the set Ω; that is, the function which is equal to 1 in points of Ω and equal to 0 elsewhere. Forming the Poisson integral

$$u(z) = \frac{1}{2\pi}\int_0^{2\pi} f(e^{it})\,\frac{1-r^2}{1+r^2-2r\cos(\theta-t)}\,dt,\ z = re^{i\theta},$$

we obtain a bounded harmonic function, not identically a constant, in the circle $|z| < 1$. This function is automorphic with respect to Γ. Indeed, since Ω is an invariant set, its characteristic function $f(e^{i\theta})$ is invariant under all transformation of Γ. But now it is easily verified that

$$u(T_n(z)) = \frac{1}{2\pi}\int_0^{2\pi} f(T_n(e^{it}))\,\frac{1-r^2}{1+r^2-2r\cos(\theta-t)}\,dt.$$

This equation and the fact that $f(T_n(e^{it})) = f(e^{it})$ imply that

$$u(T_n(z)) = u(z),\quad (n = 1, 2, \ldots),$$

which proves that $u(z)$ is automorphic with respect to Γ. The proof of the theorem is complete.

3. Let us consider a few applications of the theorem. Suppose that the group Γ has the property that its fundamental domain[4] D lies interior to a circle which is concentric with the principal circle E and has a smaller radius. It follows immediately from the theorem just proved that the group Γ is metrically transitive on E.[5] Indeed, an automorphic harmonic function which is not identically a constant could not attain its maximum in the interior or on the boundary of D. On the other hand, it would have to attain its maximum in the interior or on the boundary of D, because the function is automorphic. Hence, every automorphic harmonic function is identically a constant. Therefore Γ is metrically transitive on E.

Metrical transitivity on the principal circle may, however, also be proved for some types of Fuchsian groups whose fundamental domain extends to the circumference of the principal circle E. As an illustration of this, consider a region R_Σ in the complex w-plane formed by removing from the plane a closed bounded set of points Σ whose transfinite diameter[6] is equal to zero. Form the universal covering surface[7] of R_Σ which is known to be simply connected. By the Poincaré-Koebe theorem of uniformization there exists a single-valued meromorphic function $w = F(z)$ in $|z| < 1$ which maps the unit circle conformally on this covering surface. It is also a familiar fact of the theory of uniformization that the mapping function $F(z)$ is automorphic with respect to a certain Fuchsian group Γ which is determined by the region R_Σ. We shall prove that this group is metrically transitive on E. Suppose, indeed, that there exists a bounded harmonic function $u(z)$ in the circle $|z| < 1$ automorphic with respect to Γ. Let $z = \Phi(w)$ be the inverse function of the mapping function $w = F(z)$. Then, the function $u(\Phi(w))$ will be a bounded, single-valued, harmonic function in the region R_Σ. By a theorem of Myrberg[8] it will then be harmonic in all points of the set Σ, and consequently in all points of the w-plane. But then $u(\Phi(w))$ is identically equal to a constant. Hence, $u(z)$ is a constant. This proves by our theorem that the group Γ is metrically transitive on E.

A very special case of this result is the elliptic modular group. It may be generated in the manner described above by a region R_Σ, where Σ reduces to the three points $0, 1, \infty$. The fact that one of the points of Σ is at infinity requires an obvious modification of the argument, which shows that the elliptic modular group is metrically transitive on E.[9] Other examples of this type may be readily constructed.

4. The theorem proved in section 1 may also be used to establish conditions that a Fuchsian group be *not* metrically transitive. Let us keep the notation of the last section. We shall now assume that the set Σ is bounded, closed and such that every neighborhood of every point of Σ contains a subset of Σ with a positive transfinite diameter. If it is

assumed, then, that Σ does not decompose the w-plane, the complementary set R_Σ is a region. It was proved by Myrberg[10] that under these conditions there exists a bounded harmonic function $U(w)$, not identically a constant, in the region R_Σ for which every point of Σ is a singular point. Consider now the mapping function $w = F(z)$ of the region R_Σ on the unit circle $|z| < 1$ and form the function $U(F(z))$ which is bounded and harmonic in $|z| < 1$. Since $F(z)$ is automorphic with respect to the corresponding group Γ, $U(F(z))$ will also be automorphic. By our theorem, therefore, Γ is *not* metrically transitive on E.

Under the above assumptions on Σ there exists on E a set whose measure lies between 0 and 2π which is invariant under all transformations of the corresponding group Γ. If the further restriction is made on Σ that it shall have no connected subsets, then it may be easily shown that every open invariant set on E is everywhere dense on E. Thus the group Γ in that case would be regionally transitive[11] on E, while remaining metrically intransitive.[12]

[1] For the general theory of Fuchsian group cf. L. R. Ford, *Automorphic Functions*, Chapter III (1929).

[2] The corresponding definition in dynamics was first introduced by G. D. Birkhoff and P. A. Smith, *Jour. Mathémat.*, **7**, 9th series, 365 (1928).

[3] P. Fatou, *Acta Math.*, **30**, 366–368 (1906).

[4] Cf. L. R. Ford, *Ibid.*

[5] The metrical transitivity of these groups in phase space was proved with other methods by G. A. Hedlund, *Ann. Math.*, **35** (2), 787–808 (1934).

[6] The notion of the transfinite diameter of a point set is due to M. Fekete, *Math. Zeits.*, **17** (1923).

[7] For the general facts of the theory of uniformization cf. L. R. Ford, *Ibid.*, or H. Weyl, *Die Idee der Riemannschen Fläche* (1913).

[8] P. J. Myrberg, *Acta Math.*, **61**, 54 (1933).

[9] This fact was proved by a different method by M. H. Martin, *Bull. Am. Math. Soc.*, **40**, 606–612 (1934).

[10] P. J. Myrberg, *Ibid.*

[11] Cf. G. D. Birkhoff, *Dynamical Systems*, 205–208 (1927).

[12] After this note had been completed Professor E. Hopf informed the author that he had obtained results in a similar direction.

LARGEST GROUPS DETERMINED BY THE SQUARES OF THEIR OPERATORS

BY G. A. MILLER

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ILLINOIS

Communicated June 6, 1935

The squares of the operators of a given group G generate a characteristic subgroup H of index 2^m which includes the ϕ-subgroup of G, and H is the only invariant subgroup of G which gives rise to an abelian quotient group of order 2^m and of type $(1, 1, 1, \ldots)$. In particular, a necessary and sufficient condition that the squares of the operators of G generate exactly half of its operators is that G contains one and only one subgroup of index 2. Every group contains one and only one invariant subgroup of smallest order which gives rise to an abelian quotient group of order 2^m and of type $(1, 1, 1, \ldots)$ and this subgroup is generated by the squares of its operators. A necessary and sufficient condition that G and H are identical is that G contains no subgroup of index 2. The squares of the operators of the direct product of G and an arbitrary abelian group of order 2 and of type $(1, 1, 1, \ldots)$ also generate H. In what follows it will be assumed that G is not such a direct product.

When H is the alternating group of degree n then the largest G is the corresponding symmetric group except when $n = 6$. In this special case the order of the largest G is twice the order of the corresponding symmetric group. This statement results directly from the fact that the symmetric group of degree n is a complete group whenever $n \neq 6$ and in this case there is a complete group of order 1440 which contains this symmetric group invariantly. When G is the simple group of order 504 then it is identical with H and the largest G is H itself since the group of isomorphism of this simple group is of order 1512 as results from the theorem that when G is a transitive substitution group of degree n which involves subgroups of degree $n-1$ but no subgroups of degree n and of index n then the group of isomorphisms of G can be represented as a transitive substitution group of degree n which contains G as an invariant subgroup.[1]

Suppose that G is of order 2^n. The squares of all the operators of G are then found in every one of its subgroups of index 2 and hence in the cross-cut of all these subgroups. As this cross-cut is the ϕ-subgroup of G it results that H is then this ϕ-subgroup. Hence there results the theorem that the group generated by the squares of the operators of a group of order 2^n is its ϕ-subgroup. Whenever the order of G is not a power of 2 then the squares of its operators generate a group which includes its ϕ-subgroup but is larger than this subgroup since an operator of odd order can then appear in a set of independent generators, while all the operators of G

which correspond to operators in G/H besides the identity are of even order. Hence there results the following theorem: *A necessary and sufficient condition that the ϕ-subgroup of a group contains the squares of all of its operators is that the order of the group is a power of* 2.

When H contains an invariant operator of order 2 then the direct product of G and a subgroup of order 2 generated by a given operator s can be extended by an operator of order 2 which is commutative with every operator of G and transforms s into itself multiplied by this invariant operator. As the resulting group can be extended in a similar manner it results that there is no largest group whose squares generate H whenever H involves an operator of order 2 which is invariant under G. In particular, there is no largest group which has the property that the squares of all its operators constitute the group of order 2. To prove that there is always a largest such group when H does not contain an operator of order 2 which is invariant under G we assume first that there is no largest such group and then prove that this gives rise to a contradiction. To do this we assume that G is such a group whose order may be supposed to be as large as we choose to make it.

Since H contains no operator of order 2 which is commutative with every operator of G we may assume that G has the same property since it is not a direct product and H contains its commutator subgroup. There is a subgroup of finite order under G which transforms G in every way in which it is transformed under G. Every co-set of G with respect to this subgroup contains an operator which is commutative with every operator of G. All the operators which are commutative with every operator of G constitute a subgroup of even order under G. This subgroup therefore contains an operator of order 2 which is commutative with every operator of G. This operator does not appear in G and hence it is invariant under G so that G is a direct product of a group whose squares generate H and an abelian group of type $(1, 1, 1, \ldots)$. As this is contrary to the hypothesis it has been proved that *a necessary and sufficient condition that there is a largest group G which satisfies the condition that the squares of its operators generate a given group H is that H contains no operator of order 2 which is invariant under G.*

When H is cyclic and of odd order then the order of the largest G in which H is the group of the squares of its operators is known to be $h.2^n$, h being the order of H and n being the number of the different prime numbers which divide h. When h is even then there is no such largest group but there is a largest group such that the squares of all its operators not contained in H have for their orders the highest power of 2 which divides h. The order of this group is $2h$. When h is not a power of 2 this is the only G in which the orders of all the operators which do not appear in H have for their orders powers of 2.

From what precedes it results directly that when h is a power of 2 then there never is a largest group which has H for the group generated by the squares of its operators and that when h is odd there is always such a largest group. When h is even but not a power of 2 then there may be such a largest group, as is the case when H is the alternating group of degree $n > 2$, or there may be no such largest group, as is the case when H is the cyclic group of order 6. When H is abelian and of odd order then its holomorph is known to be a complete group[2] and no complete group can be the group generated by the squares of the operators of a largest group since this largest group would be the direct product of a group the squares of whose operators generate H and an abelian group of type $(1, 1, 1, \ldots)$.

If the squares of two operators of a group appear in a subgroup which includes their commutator then the square of their product appears in the same subgroup. From this it follows that the subgroups of the holomorph of a given group which are separately composed of operators whose squares appear in this group correspond to the largest abelian subgroups of order 2^m and of type $(1, 1, 1, \ldots)$ in the quotient group of this holomorph with respect to the given group. It therefore results that *a largest group which has the property that the squares of its operators generate a given abelian group of odd order corresponds to a largest abelian group of order* 2^m *and of type* $(1, 1, 1, \ldots)$ *in the quotient group of the holomorph of the given abelian group of odd order with respect to this group.*

In finding a necessary and sufficient condition which G must satisfy in order that there is no largest group such that the squares of its operators generate H use was made of subgroups of the final group which satisfy the condition that the squares of their operators generate H and that they are direct products of such a group and a group of order 2. If it is assumed that the final group does not contain such a subgroup then there is always such a largest group. This results from the fact that if no such largest group would exist we could find in it a subgroup K all of whose operators are commutative with every operator of G and which would involve all the operators of an indefinitely large group having this property. From the method used above it results that this subgroup would contain an operator of order 2 which is commutative with every operator of G and is not contained in G. Hence the given direct product would appear as a subgroup in the indefinitely large group. In particular, when H is of order 2 the quaternion group and the octic group are the largest groups which have the property that they do not involve a direct product of a group of order 2 and a group the squares of whose operators generate H.

[1] Miller, Blichfeldt and Dickson, *Finite Groups*, 172 (1916).

[2] G. A. Miller, *Math. Annalen*, **66**, 135 (1908).

ON UNIVERSAL HOMOLOGY GROUPS

BY NORMAN E. STEENROD

DEPARTMENT OF MATHEMATICS, PRINCETON UNIVERSITY

Communicated June 4, 1935

1. Let K be an infinite simplicial complex, and $\mathfrak{G}$ an assigned commutative group of coefficients for the chains of K. Among the cycles of K, two fundamental types appear each giving rise to its own homology theory: the finite and infinite cycles. Cech has proved[1] that the homology groups of K based on finite cycles with coefficients from $\mathfrak{J}$, the group of integers, are the universal groups of K in the sense that they determine the homology groups of K based on finite cycles with coefficients from an arbitrary $\mathfrak{G}$. Alexandroff has surmised[2] that a similar result holds for the homology theory of compact spaces, but with the group $\mathfrak{X}$ of real numbers reduced mod 1 in place of $\mathfrak{J}$. This suggests that the homology groups of K based on infinite cycles over $\mathfrak{X}$ may be the universal groups for the homology theory of the infinite cycles of K. We have recently succeeded in proving the latter surmise. The conjecture of Alexandroff follows as a corollary. The complete proof will be given elsewhere. In this note we describe the theorem and procedure, and draw some conclusions.

2. Since a discussion of the topological nature of the homology groups of the infinite cycles of K is unavoidable, we assume at the outset that $\mathfrak{G}$ is a topological group. The group of infinite p-chains over $\mathfrak{G}$ of K is topologized; and it is proved that the subgroup of cycles is closed. $\mathfrak{H}_p(\mathfrak{G})$, the pth homology group over $\mathfrak{G}$ of K, is defined as the factor group of the group of p-cycles by the *closure* of the subgroup of bounding cycles.[3]

3. K is then considered as the sum of an expanding sequence of finite subcomplexes $\{K^m\}$. Denoting by $\mathfrak{H}_p^m(\mathfrak{G})$ the pth homology group over $\mathfrak{G}$ of $K^m \mod (\overline{K-K^m})$, we define a continuous homomorphic mapping π_m of $\mathfrak{H}_p^{m+1}(\mathfrak{G})$ into $\mathfrak{H}_p^m(\mathfrak{G})$. We find that $\mathfrak{H}_p(\mathfrak{G})$ is the *limit group* of the *inverse homomorphism sequence*[4] $\{\mathfrak{H}_p^m(\mathfrak{G})\}$ in the sense that each element of $\mathfrak{H}_p(\mathfrak{G})$ is a sequence $\{x_m\}$, $x_m \in \mathfrak{H}_p^m(\mathfrak{G})$, $\pi_m(x_{m+1}) = x_m$.

4. The principal result is embodied in:

THEOREM 1. *$\mathfrak{H}_p(\mathfrak{G})$ is the direct sum of two groups, one determined uniquely by $\mathfrak{G}$ and $\mathfrak{H}_p(\mathfrak{X})$, the other by $\mathfrak{G}$ and $\mathfrak{H}_{p+1}(\mathfrak{X})$.*

$\mathfrak{H}_p(\mathfrak{G})$ is a function of the named groups in the following way. According to Pontrjagin,[5] a compact commutative group satisfying the second countability axiom may be represented as the limit group of an inverse homomorphism sequence of groups each the direct sum of a finite group and a finite dimensional toral group. Forming such representations $\{\overline{\mathfrak{H}}_p^m(\mathfrak{X})\}$, $\{\overline{\mathfrak{H}}_{p+1}^m(\mathfrak{X})\}$ of $\mathfrak{H}_p(\mathfrak{X})$ and $\mathfrak{H}_{p+1}(\mathfrak{X})$, respectively, we treat $\{\overline{\mathfrak{H}}_p^m(\mathfrak{X})\}$, $\{\overline{\mathfrak{H}}_{p+1}^m(\mathfrak{X})\}$ as if they were the groups, designated in No. 3

(with $\mathfrak{X}$ in place of $\mathfrak{G}$), for a hypothetical sequence $\{\overline{K}^m\}$. Since $\mathfrak{X}$ is a universal group for finite complexes, we can construct the homology group $\overline{\mathfrak{H}}_p^m(\mathfrak{G})$ of $\overline{K}^m$ as a function of $\mathfrak{G}$, $\overline{\mathfrak{H}}_p^m(\mathfrak{X})$, $\overline{\mathfrak{H}}_{p+1}^m(\mathfrak{X})$. In terms of the homomorphisms: $\overline{\mathfrak{H}}_p^{m+1}(\mathfrak{X})$ into $\overline{\mathfrak{H}}_p^m(\mathfrak{X})$, $\overline{\mathfrak{H}}_{p+1}^{m+1}(\mathfrak{X})$ into $\overline{\mathfrak{H}}_{p+1}^m(\mathfrak{X})$, we construct a homomorphism λ_m of $\overline{\mathfrak{H}}_p^{m+1}(\mathfrak{G})$ into $\overline{\mathfrak{H}}_p^m(\mathfrak{G})$. Then $\mathfrak{H}_p(\mathfrak{G})$ is the limit group of the inverse homomorphism sequence $\{\overline{\mathfrak{H}}_p^m(\mathfrak{G})\}$.

The proof consists in taking an actual sequence $\{K^m\}$ as in No. 3, and establishing the decomposition of $\mathfrak{H}_p(\mathfrak{G})$ into a direct sum. The sequence $\{K^m\}$ is so chosen that the related sequences $\{\mathfrak{H}_p^m(\mathfrak{X})\}$, $\{\mathfrak{H}_{p+1}^m(\mathfrak{X})\}$ interlock with subsequences of $\{\overline{\mathfrak{H}}_p^m(\mathfrak{X})\}$, $\{\overline{\mathfrak{H}}_{p+1}^m(\mathfrak{X})\}$ to form new sequences. The homomorphisms λ_m and π_m are then compared. While they may differ, they do so, by the direct sum theorem, only in an inessential fashion.

5. A compact space F in the Hilbert parallelotope may be approximated by an infinite complex K in such a way that the $(p+1)$st homology group of the infinite cycles of K is continuously isomorphic with the pth homology (Vietoris) group of F.[6] This holds for any coefficient group. It follows that $\mathfrak{X}$ is the universal coefficient group for the homology theory of compact metric spaces.

As a corollary, any property of F implied by a property of its homology groups over some group $\mathfrak{G}$ is likewise implied by some property of its homology groups over $\mathfrak{X}$. This is true of any universal coefficient group. $\mathfrak{X}$ appears to be the most convenient universal group for the following reasons.

a. It is the simplest universal group. It is easily proved that the integers, the integers mod m, the real numbers and the rational numbers are not universal groups. The group $\mathfrak{Y}$, of rational numbers reduced mod 1, may be a universal group. However homology groups are topological groups, and all operations on them are continuous. From this point of view, the group $\mathfrak{Y}$ is not simpler than $\mathfrak{X}$.

b. Freudenthal has announced [7] that the Hopf theorem on the mapping classes of the n-complex K_n on the n-sphere can be reinterpreted in terms of the character group of the nth homology group over $\mathfrak{X}$ of K_n; and, in this form, the theorem extends to the mappings of a compact metric n-space. The Hopf theorem as originally formulated does not admit of such an extension.

c. Alexandroff, Pontrjagin and Hopf have announced[8] a joint proof that the dimension modulo $\mathfrak{Y}$ (the rationals mod 1) of a compact metric space is the Menger-Urysohn dimension. As dimension is defined in terms of the non-vanishing of certain homology groups, the same theorem must hold for the dimension modulo $\mathfrak{X}$.

6. In connection with the result of Cech, we believe that the group $\mathfrak{X}$ is not a universal group for the homology theory of the *finite* cycles of an infinite K. It is not difficult to show that the pth homology group over $\mathfrak{G}$

of K where only finite cycles are considered, is the limit group of a *direct* homomorphism sequence of homology groups of finite subcomplexes. Thus, when dealing with a homology theory involving a direct homomorphism sequence, the integers form the natural coefficient group; and, when dealing with a homology theory involving an inverse homomorphism sequence, the group $\mathfrak{X}$ is the natural coefficient group. In dealing with the local homology theory of compact metric spaces, it seems to us that both of these notions are used simultaneously. The theory seems to require a direct homomorphism sequence of inverse homomorphism sequences. This complication appears to cast some doubt on the possible existence of a universal coefficient group for the theory of local homology groups.

7. The following theorem is a simple consequence of the methods used in this paper.

THEOREM 2. *Every compact, connected, commutative, topological group satisfying the 2nd countability axiom is continuously isomorphic with its own 1-dimensional homology group over* $\mathfrak{X}$.

[1] We have had the opportunity of reading this paper in manuscript form. It is to appear in *Fundamenta Mathematicae*.

[2] P. Alexandroff, "On Local Properties of Closed Sets," *Ann. Math.*, **36**, 1–35 (1935).

[3] Cf. S. Lefschetz, "Topology," *Amer. Math. Soc. Pub.*, **12**, 299 (1930).

[4] This notion was developed by L. Pontrjagin, "Über den algebraischen Inhalt topologischer Dualitätssätze," *Math. Ann.*, **105**, 165–205 (1931). C. Chevalley has remarked that the limit group of such a sequence may be used in defining the Vietoris group of a compact space; *Compt. Rend.*, **200**, 1005–1007 (1935).

[5] L. Pontrjagin, "The Theory of Topological Commutative Groups," *Ann. Math.*, **35**, 361–388 (1934).

[6] *Topology*, pp. 323–334.

[7] H. Freudenthal, "Die Hopfsche Gruppe," *Compositio Math.*, **2**, 134–162 footnotes 2a and 8 (1935).

[8] P. Alexandroff, l. c. (2), p. 27.

A CONFORMAL WAVE EQUATION

BY OSWALD VEBLEN

THE INSTITUTE FOR ADVANCED STUDY

Communicated June 6, 1935

The subject of this note is a differential equation in generalized four-dimensional conformal geometry, which is somewhat analogous to the Dirac equation for the electron in classical relativity. It is presumably equivalent in case of the classical conformal geometry to an equation which has been developed independently by Dirac in a different notation.[1]

We use the notation outlined in a recent note in these PROCEEDINGS.[2]

The fundamental conformal tensor is[3] $g_{\sigma\tau}$ and it determines a conformal connection $C^{\sigma}_{\tau i}$ in terms of which a conformal derivative is defined which has an affine index. We shall convert affine indices into alternating pairs of conformal ones and vice versa by Vanderslice's method, using

$$A^{i}_{\sigma\tau} = \frac{1}{2}(\delta^{i}_{\sigma}\delta^{5}_{\tau} - \delta^{i}_{\tau}\delta^{5}_{\sigma}). \tag{1}$$

Conformal tensors may be converted into four-component spinors, and spinors into conformal tensors, by means of spinors g^{σ}_{AB} and $g^{\sigma AB}$ which are antisymmetric in the spin indices AB, and so related that

$$g^{\sigma AB} = \frac{1}{2}\epsilon^{ABCD}g^{\sigma}_{CD}. \tag{2}$$

In accordance with (2) we assume that g^{σ}_{AB} is of weight $-1/2$, and $g^{\sigma AB}$ of weight $+1/2$. Geometrically these spinors define a Pluecker-Klein correspondence

$$X^{\sigma} = g^{\sigma}_{AB}X^{AB} \tag{3}$$

between the points X^{σ} of the real four-dimensional conformal tangent space, and certain straight lines X^{AB} of a complex three-dimensional spin-space. When using a matrix notation for spinors, we denote g^{σ}_{AB} by g^{σ} and $g^{\sigma AB}$ by $g^{\sigma\sim}$. These spinors satisfy the relation

$$g^{\sigma\sim}g^{\tau} + g^{\tau\sim}g^{\sigma} = -\frac{1}{2}g^{\sigma\tau}1, \tag{4}$$

where $g^{\sigma\tau}$ is the fundamental conformal tensor.

The spinor g^{σ} also satisfies the relation[4]

$$\bar{g}^{\sigma}_{\dot{A}B} = g_{\dot{A}C}g_{\dot{B}D}g^{\sigma CD} \tag{5}$$

in which $g_{\dot{A}C}$ is an anti-Hermitian spinor of signature $(++--)$ and the bar is used to indicate the complex conjugate. The components of this spinor are the coefficients of the equation of an antiquadric,

$$g_{\dot{A}B}\bar{X}^{A}X^{B} = 0 \tag{6}$$

in the complex three-dimensional projective spin-space. The ∞^4 lines on this antiquadric are represented by the antisymmetric spinors X^{AB} which are related to real conformal vectors X^{σ} by (3). The signature of $g_{\dot{A}B}$ is chosen so that the tangent conformal space shall have the right signature for the relativity theory.

Covariant differentiation of spinors is defined in the usual way in terms of a spin connection Γ^{A}_{Bi}. Imposing the condition,[5] $g^{\sigma}_{;i} = 0$, we find that the traceless part of this spin connection is determined; that is to say, if

$$K_i = \Gamma_i - \frac{1}{4}(\text{Trace } \Gamma_i)1 \tag{7}$$

then

$$K_i = -\left(g_\sigma^{\sim} \frac{\partial g^\sigma}{\partial x^i} + C^\sigma_{\tau i} g_\sigma^{\sim} g^\tau\right). \tag{8}$$

When Γ_i satisfies this condition it follows that

$$g_{\dot{A}B,i} = 0. \tag{9}$$

We thus have Γ_i in the form

$$\Gamma_i = K_i + \frac{1}{4}\varphi_i 1 \tag{10}$$

in which the functions φ_i are arbitrary but have the law of transformation of the components of a covariant affine vector under transformations of coördinates and the law of transformation

$$\varphi_i^* = \varphi_i + \frac{\partial \log |t|}{\partial x^i} \tag{11}$$

under spin transformations.

If we restrict our spin transformations so that their determinant is of absolute value 1, i.e., so that

$$|t| = e^{i\theta} \tag{12}$$

where θ is real and if we set $\varphi_i = \sqrt{-1}A_i$, (11) becomes

$$A_i^* = A_i + \frac{\partial \theta}{\partial x^i}. \tag{13}$$

Thus A_i behave like the components of an electromagnetic vector, the gauge change being determined by the spin transformation.

For an arbitrary simple spinor ψ of weight w, the formula for covariant differentiation is now

$$\psi_{,j} = \frac{\partial \psi}{\partial x^j} + K_j\psi + i\left(\frac{1}{4} - w\right)A_j\psi. \tag{14}$$

The wave equation which we are seeking should be a linear combination of these covariant derivatives with appropriate matrices as coefficients.

The fifteen-fold totality of matrices of the type which appear in the Dirac equation are embodied in the spinor,

$$s^{\sigma\tau} = \frac{1}{2}(g^\sigma{}^{\sim} g^\tau - g^\tau{}^{\sim} g^\sigma). \tag{15}$$

The Vanderslice method of converting alternating pairs of conformal indices into affine ones suggests that we form

$$s^j = A^j_{\sigma\tau} s^{\sigma\tau} = \frac{1}{2}(g^j{}^\sim g^5 - g^5{}^\sim g^j) = s^{5j} \tag{16}$$

and then have the wave equation

$$s^j \psi_{,j} = M\psi \tag{17}$$

where M could be a spinor. If M is taken to be a constant m, we have the equation

$$s^j \left(\frac{\partial \psi}{\partial x^j} + K_j \psi + aiA_j \psi \right) = m\psi \tag{18}$$

in which a and m are constants.

The matrices s^j are of rank two whereas those which appear in the regular Dirac equation are of rank four. Geometrically, this is because, for fixed values of σ and τ the matrix $s^{\sigma\tau}$ represents a line reflection in the projective spin-space and this line reflection is imaged by a line in P_5. The line reflection degenerates whenever the line in P_5 is tangent to the fundamental quadric and this happens in case of the matrices s^{5j}. In the case of the ordinary Dirac equation the four matrices γ^i represent lines through a point not on the quadric. The invariance of this point reduces the group relevant to this equation from the conformal group to the projective relativity group.

[1] P. A. M. Dirac, "Wave Equations in Conformal Space," to appear in the *Ann. Math.*, **37** (1936).

[2] These PROCEEDINGS, **21**, 168–173 (1935).

[3] Small Roman letters are used for affine indices and run over the range 1, ..., 4. Small Greek letters are used for conformal indices and run over the range 0, ..., 5. (We used ∞ in the previous note for the index here denoted by 5.) Large Roman letters are used for spin indices, and run over the range 1, ..., 4.

[4] Cf. "Geometry of Four-Component Spinors," these PROCEEDINGS, **19**, 503–517 (1933).

[5] For these results, including Equations (7), (8), (9), I am indebted to Mr. J. W. Givens.

THE KINETICS OF PENETRATION. X. GUANIDINE

BY A. G. JACQUES

LABORATORIES OF THE ROCKEFELLER INSTITUTE FOR MEDICAL RESEARCH, AND THE BERMUDA BIOLOGICAL STATION FOR RESEARCH, BERMUDA

Communicated May 29, 1935

In a former paper[1] from this Laboratory it was shown that the rate of entrance of the weak base ammonia into the large multinucleate cell of *Valonia macrophysa* Kütz fails to increase as rapidly as the concentration gradient of molecular $NH_3 + NH_4OH$, although it should do so if the mode of entrance merely involved the diffusion of these substances through a non-aqueous layer.

To account for this discrepancy a simple explanation has been offered according to which the entrance of the base is preceded by a reversible reaction

$$NH_4OH + HX \rightleftharpoons NH_4X + H_2O$$

where HX is an unidentified acidic constituent of the protoplasm.

The rate is then assumed to be proportional to the concentration gradient of NH_4X in the non-aqueous layer, and if HX is scarce in amount as compared with $NH_3 + NH_4OH$ the decrease in rate with increasing concentration of the last two is inevitable.

The question now arises whether a similar process is involved in the entrance of strong bases.

In the case of *Nitella flexilis* it has already been shown[2] that the rate of entrance of potassium reaches a maximum beyond which a further increase in the external concentration of potassium fails to increase the rate. But this is not a clear-cut case, for the internal concentration of potassium is, at the start, much greater than the external concentration, and, as far as we can tell, at present, all the concentration gradients, between the sap and the surrounding medium suggest that potassium should leave the sap.

As a less involved case, it was decided to study the rate of entrance of a strong base, not already present in the sap, into *Valonia*. Of the strong inorganic bases, lithium and rubidium are known to enter the cell.[3] The analytical procedures for small amounts of these bases are, however, somewhat tedious and inaccurate, and the choice therefore fell on guanidine, $NH_2 \diagdown C = NH$, an organic base which with some of its simpler aliphatic $NH_2 \diagup$ derivatives is known to be as strong a base as sodium hydroxide.[4]

The experiments were carried out in Bermuda during the winter of 1934–35. They were performed in subdued light so that the external pH was not affected seriously by photosynthesis.

Experimental.—The sea-water samples containing guanidine were prepared by adding to normal sea water appropriate quantities of 0.6 *M* guanidine hydrochloride solution. This was prepared from a more concentrated solution of guanidine carbonate, by adding 2 *M* hydrochloric acid until the solution was distinctly acid. During the addition the solution was exhausted and shaken to remove the CO_2. After the gas had been removed the volume of the solution was made up to yield a solution of 0.6 *M* guanidine, by adding water and 0.1 *M* sodium hydroxide to bring the pH to 8.2, that of the normal sea water.

The cells, fairly uniform in size (0.2 to 0.4 ml. in volume) were selected from collections made three weeks to a month before use. They were kept during the seasoning period in large aquaria, with flowing sea water.

The exposures were made in 125-cc. bottles, 15 to 20 cells to a bottle. It was found necessary to stir in order to get uniform results and this was achieved by fixing the bottles with large rubber bands to the outer ends of the spokes of a bicycle wheel, which was turned at the rate of 1.5 revolutions per minute by means of a motor and two speed reducing gears in series.

Analyses.—It had been hoped to use the color reaction worked out by Marston[5] but it was found that the amount of guanidine entering the cell in a reasonable time was too small to permit accurate analysis, especially in the presence of the high salt concentration of the sap. Accordingly the guanidine was determined as nitrogen by subjecting the sap to the Kjeldahl treatment, with the sulfuric-phosphate-potassium persulfate mixture according to the method of Van Slyke.[6] It was found necessary to heat the Kjeldahl mixture for at least 12 hours over an electric sand bath in order to bring about complete decomposition of the guanidine.

The presence of other nitrogenous substances in the sap was an unavoidable complication and a blank correction had to be made. However, it was found that when the sap was energetically centrifugalized to remove chloroplasts and other debris, the amount of ammonia present after Kjeldahlization was not much greater than that present in the sap before this treatment, so that the correction practically amounts to the free ammonia normally in the sap, and this in our cells which had been seasoned in running sea water was considerably less than that previously reported for cells which were seasoned in the laboratory in still sea water.[7]

It was also necessary to establish the fact that the nitrogen found in the cells was not derived from ammonia, produced by hydrolysis of the guanidine in the sea water, or from ammonium salt occurring as an impurity of the guanidine. This was done by determining directly the ammonia in the sap after prolonged exposure to a guanidine containing sea water. The amount was not appreciably greater than the amount of ammonia normally present in the cells.

After Kjeldahlization the ammonia was distilled off in a current of steam, using an apparatus similar to that described by Teorell,[8] into an excess of dilute sulfuric acid, and was determined finally by means of the Nessler reaction,[9] using a Zeiss Pulfrich "step photometer." The sap was extracted by means of a hypodermic syringe and centrifugalized for 10 to 20 minutes in a closed tube. 1 or 2 cc. samples were used for analysis.

The concentration of the guanidine in the sea water varied from 0.005 *M* to 0.05 *M*. When the concentration was raised to 0.1 *M* it was found that after 24 hours some of the cells appeared to be injured although no sulfate was detected in the sap. Accordingly the maximum concentration used in the actual experimental work was about 0.05 *M* and the exposures did not exceed 10 hours.

There was no sign of injury as shown by the appearance, turgidity and sulfate test. After the experiment the cells lived indefinitely when returned to sea water.

The temperature varied from 15 to 20°C. but did not change more than 2°C. in the course of a single experiment.

Results.—Table 1 gives for three typical experiments the concentration of guanidine found in the sap after 10 hours for cells which have been exposed to various concentrations of guanidine in sea water at pH 8.2 to 8.3.

TABLE 1

MOLAR CONCENTRATION OF GUANIDINE IN THE SAP

CONCENTRATION OF GUANIDINE IN SEA WATER *M*	EXPERIMENT 1 *M*	EXPERIMENT 2 *M*	EXPERIMENT 3 *M*
0.0050	0.00065	0.00043	0.00040
0.0100	0.00110	0.00072	0.00126
0.0175	0.00131	0.00102	.....
0.0250	0.00174	0.00155	0.00207
0.0375	0.00198	0.00193	.....
0.0500	0.00207	0.00205	0.00278

Discussion of Results.—In drawing the curve, the external concentrations of guanidine have been plotted as abscissae. Strictly, of course, the gradients should have been so plotted, and, on the assumption that guanidine enters as guanidonium hydroxide, the gradients $(G)_o(OH)_o - (G)_i(OH)_i$. A linear relationship in this case would have indicated that the mode of entrance was diffusion. However, it should be observed that the internal pH of the sap was not changed detectably during the penetration of the guanidine and hence $(OH)_o$ was roughly 100 times $(OH)_i$. Hence at the point where the experiments were terminated $(G)_i(OH)_i$ must have been very small in comparison with $(G)_o(OH)_o$ and could be neglected. Even if we make the assumption that guanidine enters in ionic form and diffuses through the protoplasm as such, the external concentration of guanidine during the period of the experiments was not much different from the gradient since in the worst case the internal concentration was only 13 per cent of the external concentration. Correction for this factor, if it were possible, would not alter much the shape of the curve except to make it a little steeper at the left.

The fact that there is a distinct curvature is explainable, as in the case of ammonia, on the assumption that the entrance of the base is preceded by a reaction

$$GOH + HX \rightleftharpoons GX + H_2O$$

The question whether the same acidic constituent HX is involved in both ammonia and guanidine penetration ought probably to be answered in the affirmative, but it is difficult to prove it by a comparison of the slope of the curves because the curvature depends on several factors, namely: the concentration of HX which might be the same always, but probably differs in different lots of cells, and perhaps even from cell to cell,

so that even in one particular lot it is an average value; the value of k, the velocity constant of the reversible reaction; the length of exposure, size of the cell and other factors which determine the limit which the rate tends to approach.

Attempts were made to fit the data to the equation

$$G_o[(HX_b) - (GX_e)] = k(GX_e)$$

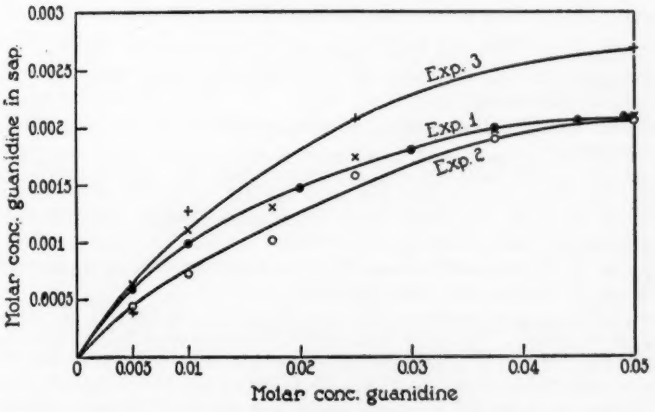


FIGURE 1

Rate of penetration of guanidine into *Valonia macrophysa* in 10 hours at pH 8.2 to 8.3 at various external guanidine concentrations.

The solid circles represent the points calculated for the equation

$$G_o\,[(HX_b) - (GX_e)] = k(GX_e)$$

where G_o is the concentration of guanidine in the sea water. HX_b is the concentration of the acidic constituents of the protoplasm at the beginning of the reversible reaction

$$GOH + HX \rightleftharpoons GX + H_2O$$

and GX_e is the concentration of the new salt at the equilibrium point of the reaction. (HX_b) is taken as 0.00295 and $k = 0.02$. The curves are drawn free-hand to give a rough fit.

The other points represent observations in Experiments 1, 2 and 3.

where G_o is the concentration of guanidine, (HX_b) the concentration of the acidic carrier at the beginning of the reaction and (GX_e) the concentration of the new molecular species at the equilibrium of the reversible reaction. Since the rate is assumed to be proportional to (GX_e) we may substitute the latter for it. In the experiments in question the "rate" is taken as the amount of guanidine which has entered the cell in 10 hours. By putting $(HX_b) = 0.00295$ and $k = 0.2$ it was found that the data of Experiment 1 (table 1) fitted very well. However, neither Experiment 2

nor 3 could be fitted so closely to this type of equation; nor could the average of the three experiments. For these reasons we prefer not to emphasize the fit in the one case.

Summary.—When the rate of entrance of guanidine is plotted against the external concentration of guanidine, a curve which is concave to the horizontal axis is obtained. This is explainable on the assumption that the entrance is preceded by a reversible action between the penetrating base and one or more acidic constituents of the protoplasm.

[1] Osterhout, W. J. V., *Proc. Nat. Acad. Sci.*, **21**, 127 (1935).

[2] Jacques, A. G., and Osterhout, W. J. V., *Proc. Soc. Exptl. Biol. Med.*, **31**, 1121 (1933–34).

[3] Cooper, W. C., Jr., Dorcas, M. J., and Osterhout, W. J. V., *Jour. Gen. Physiol.*, **12**, 427 (1928–29).

[4] Davis, T. L., and Elderfield, R. C., *Jour. Am. Chem. Soc.*, **54**, 1499 (1932).

[5] Marston, H. R., *Austral. Jour. Exp. Biol. and Med. Sci.*, **1**, 99 (1924); **2**, 57 (1925).

[6] Peters, J. P., and Van Slyke, D. D., *Quantitative Clinical Chemistry*, Baltimore, Williams and Wilkins, **2**, 23 (1932).

[7] Cooper, W. C., Jr., and Osterhout, W. J. V., *Jour. Gen. Physiol.*, **14**, 117 (1930–31).

[8] Teorell, T., *Biochem. Zeit.*, **248**, 246 (1932).

[9] Nessler's reagent according to Folin and Wu (Folin, O., and Wu, H., *Jour. Biol. Chem.*, **38**, 81 (1919)). Two drops of gum ghatti was added to each 50 cc. of the mixture before the addition of the Nessler reagent according to the advice of Folin and Svedberg (Folin, O., and Svedberg, A., *Jour. Biol. Chem.*, **88**, 77 (1930)). Urbach (Urbach, C., *Mikrochemie*, **11**, 37 (1932)) has shown that a linear relationship between the extended coefficient and concentration with the purple filter of the Zeiss photometer can be obtained under the best conditions. In our work the calibration curve was not quite linear, but had so little curvature that no serious error was introduced.

THE CORRELATION BETWEEN EXCESS CALORIES AND EXCESS URINARY NITROGEN IN THE SPECIFIC DYNAMIC ACTION OF PROTEIN IN ANIMALS

BY HENRY BORSOOK

WILLIAM G. KERCKHOFF LABORATORIES OF THE BIOLOGICAL SCIENCES, CALIFORNIA INSTITUTE OF TECHNOLOGY

Communicated June 10, 1935

Several years ago it was pointed out that after the ingestion of proteins or amino acids a definite correlation existed between the calories and urinary nitrogen in excess of the basal.[1] The data used were those obtained by Rapport,[2] Weiss and Rapport[3] and Rapport and Beard.[4] Aubel and Schaeffer[5] were of the opinion that this correlation bears no relation to the specific dynamic action of protein and amino acids. Their main reason was that the curves of extra calories and extra nitrogen are not

synchronous; that the curve of caloric metabolism returns to the basal level before that of the nitrogen. They maintained that no correlation is to be found between excess calories and excess nitrogen in any one experiment; and cited data from the experiments of Rapport and Beard[4] which they maintained show that very large differences in specific dynamic action of individual amino acids occur which bear no relation to the extra nitrogen excreted.

This criticism is rooted in the refusal of Aubel and Schaeffer, in accord with the views of the Lusk School, to ascribe any significant fraction of the specific dynamic action to the heat evolved in the metabolism and excretion of the nitrogen. Yet both Lusk and Aubel agreed that the specific dynamic action of an amino acid is to be correlated with the amount metabolized. The most obvious measure of this quantity is the urinary nitrogen. Because the nitrogen excretion lags behind the increase in metabolism Lusk preferred to use the extra glucose excreted in a phlorhizinized animal. The choice between these two indices of the quantity of amino acid metabolized rests on the accuracy with which they may represent this quantity, and it is quite immaterial if there is any increased expenditure of energy in the metabolism of the nitrogen or not.

The use of the extra glucose stands and falls on the assumption that the rate of metabolism is the same in the phlorhizinized and normal animal. Examination of the data on nitrogen excretion in the experiments of Lusk[6] and of Csonka[7] which form the basis of Lusk's method show that the rate of nitrogen metabolism was not the same in the two conditions. Furthermore the increase in metabolism lagged behind the glucose excretion. This method breaks down entirely with ketogenic amino acids such as phenylalanine and tyrosine which yield no glucose and yet exert a larger specific dynamic action than glucose forming amino acids.

The nitrogen excretion is obviously a more direct index. The difficulty that the nitrogen excretion lags behind the increase in metabolism can be circumvented by giving most weight to the data during the approximate steady state when the nitrogen is excreted as fast as it is metabolized. It was for this reason that the increase in heat production (or oxygen consumption) and nitrogen excretion during the first four to six hours were used, and indeed could be used, to show the correlation between increase in calories and nitrogen excretion. The curves of energy metabolism and nitrogen excretion are not widely divergent; they cover nearly the same area (with appropriate units) and the quantity of nitrogen whose excretion is really delayed represents a small fraction of the total quantity excreted by the time the descending arm of the nitrogen curve begins to flatten out. It appears to have been tacitly assumed by Lusk and by Aubel that all of the ingested nitrogen, being exogenous, is quickly metabolized. This is not the case. In man in nitrogen balance at a moderate

TABLE 1

THE STEADY STATE BETWEEN EXCESS CALORIC METABOLISM AND URINARY NITROGEN EXCRETION AFTER THE INGESTION OF PROTEIN

HOURS	EXCESS CALORIES PER CENT OF BASAL	EXCESS NITROGEN PER CENT OF BASAL	EXCESS CALORIES EXCESS NITROGEN % OF BASAL × 100	OBSERVER
1	61	128	48	
2	87	385	23	
3	85	465	18	
4	79	495	16	
5	82	525	16	
6	98	505	19	
7	79	500	16	
8	89	500	17	Dog, 1200 gm. meat.
9	83	500	16	Williams, Riche and
10	71	450	16	Lusk[9]
11	69	450	15	
12	61	450	14	
13	65	487	13	
14	31	400	8	
15	38	400	10	
16	42	304	14	
17	39	238	16	
18	17	194	9	
19	23	141	16	
20	12	100	12	
1	11.7	85	13	
2	21.5	190	11	
3	26.5	255	9	
4	19.8	330	6	Man, 69 gm. gelatine.
5	19.6	315	6	Borsook and Keigh-
6	9.6	290	3	ley[10]
7	11.8	214	6	
8	9.6	190	5	
9	11.6	206	6	
10	13.6	170	8	
1	17.5	0	∞	
2	16.9	37	46	
3	27.6	76	36	Man, 630 gm. meat.
4	16.3	85	19	Borsook and Keigh-
5	17.8	130	14	ley[10]
6	23.8	137	17	
7	28.4	180	16	
8	23.0	142	17	

level of nitrogen intake the urinary nitrogen excreted in 24 hours is not more than fifty per cent of exogenous origin.[8]

Table 1 contains a summary of one experiment of Williams, Riche and Lusk[9] and of two typical examples of our own[10] showing the attainment of a steady relation between increase in energy metabolism and urinary nitrogen excretion. This steady state in our experiments appeared in most cases after the fourth hour.

The same phenomenon is shown in another form in table 2. This form permits (on account of space limitations) the presentation of more experiments than table 1. The ratio of excess calories to excess nitrogen is, except where a very large quantity of meat was ingested, nearly the same over the first four hours as over a much longer period. Even though the

TABLE 2

RELATION BETWEEN EXCESS CALORIES AND URINARY NITROGEN IN EXCESS OF THE BASAL AFTER THE INGESTION OF PROTEIN

SUBJECT OF EXPERIMENT; SUBSTANCE AND QUANTITY INGESTED	OBSERVER	TOTAL PERIOD OF OBSERVATION, HRS.	EXCESS CALORIES EXCESS NITROGEN FOR TOTAL PERIOD	EXCESS CALORIES EXCESS NITROGEN FOR FIRST FOUR HRS.
Dog, 1200 gm. meat	Williams, Riche & Lusk[9]	20	11.0	10
Dog, 700 gm. meat	Williams, Riche & Lusk[9]	7	13.0	13
Man, 69 gm. gelatine	Borsook & Keighley[10]	10	31	30
Man, 69 gm. gelatine	Borsook & Keighley[10]	10	18	16
Man, 69 gm. gelatine	Borsook & Keighley[10]	10	18	21
Man, 69 gm. gelatine	Borsook & Keighley[10]	10	20	24
Man, 32.8 gm. gelatine	Borsook & Keighley[10]	10	28	31
Man, 315 gm. meat	Borsook & Keighley[10]	10	38	38
Man, 315 gm. meat	Borsook & Keighley[10]	10	18	25
Man, 630 gm. meat	Borsook & Keighley[10]	10	21	45
Man, 150 gm. meat	Borsook & Keighley[10]	10	29	21
Man, 315 gm. meat	Borsook & Keighley[10]	10	25	27
Man, 450 gm. meat	Borsook & Keighley[10]	10	28	38

nitrogen excretion lags behind the energy metabolism in the first two hours it represents only a small quantity of nitrogen. Hence the ratio is not seriously disturbed by this lag.

Data suitable for such a compilation as in tables 1 and 2 is limited because relatively few experiments have been reported which were carried out over a sufficiently long period to make such comparisons possible; in even fewer cases are the hourly urinary nitrogen as well as metabolic data given.

Aubel and Schaeffer, in criticizing the use of four to six hour averages in our original demonstration of the correlation between excess calories and excess nitrogen apparently attach more weight to the hourly data in individual experiments. Our reason for preferring averages to indi-

vidual hourly values was that the hour to hour caloric values are as irregular as the urinary nitrogen excretion. It is obvious, however, that if the averages from different experiments group themselves about a smooth curve the individual values from which these averages were obtained will also do so. It was pointed out that the deviations of any individual value from the smooth curve are within the extremes of variation of the basal metabolism values.

It is this irregularity in the metabolic data that makes it extremely difficult in the terminal stages of an experiment to determine the increase in metabolism to be attributed solely to increased renal work. At this time the increased excretion of nitrogen is not, expressed in absolute quantities, greatly in excess of the basal. Any possible increases in metabolism are well within the variations in the basal metabolism values.

Aubel and Schaeffer referred to the observations of Gigon,[11] and stated that the energy metabolism had returned to the basal level long before the nitrogen excretion. Examination of Gigon's data shows that this is

TABLE 3

EXCESS CALORIES AND URINARY NITROGEN AFTER THE INGESTION OF VARYING AMOUNTS OF CASEIN[11]

AMOUNT OF CASEIN INGESTED, GM.	PERIOD OF OBSERVATION, HRS.	AMOUNT OF CASEIN N RECOVERED IN URINE, PER CENT	EXCESS CALORIES, CAL.	EXCESS NITROGEN, GM.	EXCESS CALORIES / EXCESS NITROGEN
50 (6.4 g. N)	3.5	17	19	1.1	17
100 (12.8 g. N)	5.75	30	53	3.8	14
150 (19.2 g. N)	7	30	118	5.7	21
200 (25.6 g. N)	10	35	171	9.0	19

not the case. In the experiment where 50 gm. of casein were ingested, which Aubel and Schaeffer selected for citation, the nitrogen excretion from the fifth to the tenth hours was below the basal, 2.5 gm. as against 2.95 gm. Table 3 is a summary of the data in Gigon's experiments, with varying amounts of casein. The ratio of excess calories to excess nitrogen is not widely different through a fourfold variation in the amount of protein ingested, and a threefold difference in the period of observation. These data confirm the usefulness of the correlation of excess calories to excess nitrogen.

Aubel and Schaeffer referred to the data of Rapport and Beard which they maintained showed no relation between the specific dynamic action and the excretion of nitrogen. The data of Rapport and Beard are summarized in table 4. In interpreting these data the varying mode of metabolism of the individual amino acids must be taken into account. The specific dynamic action is a composite of two quotas, one nearly constant representing the metabolism and excretion of the nitrogen, and another

variable quota, the metabolism of the deaminized residues.[8] When nitrogen is released by hydrolytic processes the energy change will be less than in the case of oxidative deamination. In the case of arginine, urea arises directly by hydrolysis; hence the increased oxygen consumption attending deamination and urea synthesis will not occur here. With histidine ammonia arises by splitting of the imidazole ring; the energy of oxidative deamination again will be missing. The metabolism of the deaminized residues in the case of tyrosine and phenylalanine, splitting of the benzene ring, etc., will lead to a greater specific dynamic action than with alanine or glycine. The same applies to glutamic and aspartic

TABLE 4

RELATION BETWEEN EXCESS NITROGEN AND EXCESS CALORIES AFTER THE ADMINISTRATION OF DIFFERENT AMINO ACIDS TO DOGS[4]

AMINO ACID KIND	AMOUNT, GM.	NITROGEN EXCRETION IN EXCESS OF BASAL PER HOUR, GM.	CALCULATED INCREASE IN METABOLISM DUE TO WORK OF EXCRETION OF N, 5 CAL./GM. OF N CALORIES	OBSERVED INCREASE IN METABOLISM CALORIES	FRACTION OF N ADMINISTERED RECOVERED IN URINE, PER CENT
Glutamic acid	10	0.048	0.24	1.55	16
	10	0.044	0.22	1.22	16
	10	0.027	0.14	1.36	16
	20	0.033	0.17	1.32	6
Aspartic acid	10	0.032	0.16	1.09	11
	10	0.028	0.14	0.93	11
	20	0.043	0.21	1.17	6
Histidine dichloride	5	0.034	0.17	0.27	15
Cystine	10	0.024	0.12	1.23	7
Valine	10	0.11	0.55	0.47	21
Tyrosine	10	0.037	0.19	2.7	19
	10	0.033	0.16	2.1	17
Phenylalanine	10	0.10	0.51	5.0	47
	8.5	0.11	0.57	4.6	62
Arginine carbonate	5	0.038	0.19	0.65	12

acids. The data show that though Rapport and Beard, and Aubel and Schaeffer assigned no specific dynamic action to histidine and valine, the increases in metabolism observed were in every case sufficient to account for the cost of excreting the nitrogen; and in every case except valine the increased metabolism was much greater than this figure. Until more is known of the intermediary metabolism of valine it is unprofitable to speculate on the small increase in metabolism observed. Small as it was, this increase was sufficient for the excretion of the nitrogen. Further, in the case of valine, the observations were terminated at a time (the third hour) when the increase in metabolism was at its peak for the period, 0.99 calories per hour above the basal. It seems that more extended observa-

tions may have given a value for the ratio of excess calories to excess nitrogen more in accord with those of other amino acids.

Summary.—1. When protein is ingested a steady state is attained after four hours between the increase in energy metabolism and urinary nitrogen.

2. The ratio of total excess calories to total excess nitrogen is nearly the same after four hours as after ten to twenty hours.

3. These results confirm the usefulness and significance of the correlation previously demonstrated between excess calories and excess nitrogen in the specific dynamic action of protein in animals.

4. Analysis of the data upon which Aubel and Schaeffer based their criticism of the significance of this correlation shows that this criticism is invalid.

[1] Borsook, H., and Winegarden, H. M., *Proc. Nat. Acad. Sci.*, **17**, 75 (1930).
[2] Rapport, D., *Jour. Biol. Chem.*, **60**, 497 (1924).
[3] Weiss, R., and Rapport, D., *Ibid.*, **60**, 513 (1924).
[4] Rapport, D., and Beard, H. H., *Ibid.*, **73**, 299 (1927); **80**, 413 (1928).
[5] Aubel, E., and Schaeffer, G., *Ann. Physiol. Physicochim. biol.*, **8**, 262 (1932).
[6] Lusk, G., *Jour. Biol. Chem.*, **20**, 555 (1915).
[7] Csonka, F. A., *Ibid.*, **20**, 539 (1915).
[8] Borsook, H., and Keighley, G., *Proc. Nat. Acad. Sci.*, **20**, 179 (1934).
[9] Williams, H. B., Riche, J. A., and Lusk, G., *Jour. Biol. Chem.*, **12**, 349 (1912).
[10] Borsook, H., and Keighley, G., Unpublished experiments.
[11] Gigon, A., *Archiv. ges. Physiol.*, **40**, 509 (1911).

RELATION OF THE REGULATORY MECHANISM OF RESPIRATION TO CLINICAL DYSPNEA

BY DICKINSON W. RICHARDS, JR., ANDRÉ COURNAND AND ISRAEL RAPPAPORT

DEPARTMENT OF MEDICINE, COLLEGE OF PHYSICIANS AND SURGEONS, COLUMBIA UNIVERSITY, THE PRESBYTERIAN HOSPITAL, AND THE TUBERCULOSIS DIVISION OF BELLEVUE HOSPITAL, NEW YORK CITY

Communicated May 29, 1935

According to classical theory, respiration is regulated by impulses which are automatically generated within and which emanate from the "respiratory center" situated in the medulla. Recent work (Lumsden,[1] Hess,[2] Heymans,[3] Gesell[4]) has shown that this concept must be broadened, and that the regulation of respiration is a mechanism of wide extent, mediated through many "centers," in brain, spinal cord, carotid sinuses, sympathetic and para-sympathetic nervous systems, as well as in the "respiratory center" in the medulla. Both somatic and vegetative periph-

eral nervous systems—particularly the latter—are concerned in conveying the sensory and motor impulses which effect the act of breathing.

Hess,[2] in his monograph "Die Regulierung der Atmung," was led to separate the physicochemical, metabolic forces which, acting through the respiratory centers, necessitate a given volume of gaseous exchange, from the mechanical apparatus of breathing, which provides the ventilation, or form of pulmonary activity by which the gaseous exchange is accomplished. In a more recent review Fleisch[5] also conceives of the central regulating mechanism ("respiratory center") as imposing a definite ventilatory air exchange, with inspiratory and expiratory phases. The proprioceptive nervous impulses, streaming into the respiratory centers from all parts of the breathing mechanism, determine the manner and the intensity of the breathing act whereby the ventilatory requirement is carried out.

Clinically, dyspnea is a sensation of subjective distress referred by the individual to the act of breathing, or any part of it. Dyspnea does not appear at any definite level of ventilatory volume, of acidity of the blood or of oxygen unsaturation; it is, in general, not determined by any physicochemical bodily function, or group of functions, as far as these are known. Illustrations of this fact are given in the accompanying table:

TABLE 1

CONDITION	OXYGEN CONSUMPTION CC./MIN.	ARTERIAL BLOOD			PULMONARY VENTILATION LITERS/MIN.	DEGREE OF DYSPNEA
		OXYGEN SAT'N %	CO_2 TENSION	pH_s		
1. Normal, A. A. McL. exercising	2900	..	40.3	7.37	58	slight
1. Normal, W. O. exercising	2700	..	33.5	7.29	82	marked
2. Normal in experimental acidosis, rest	...	..	26.9	7.20	8.5	none
2. Same, exercising	2750	..	31.1	7.07	80	extreme
3. Diabetic acidosis at rest	200	95	(28.1)*	...	7.9	none
4. Congenital heart disease, at rest	188	60	32.0	7.48	9.4	none
5. Congestive heart failure, at rest	185	91	36.0	7.43	7.7	moderate
6. Pulmonary fibrosis, mild exercise	383	94	41.0	7.40	9.1	extreme

* This figure is CO_2 content, in volumes per cent.

1. Dill, D. B., Talbott, J. H., and Edwards, H. T., *J. Physiol.*, **69**, 18 (1930).
2. Dennig, H., Talbott, J. H., Edwards, H. T., and Dill, D. B., *J. Cl. Inv.*, **9**, 601 (1931).
3. Jansen, K., Knipping, H. W., and Stromberger, K., *Beitr. Kl. Tuberk.*, **86**, 304 (1932).
4. Richards, D. W., *Arch. Int. Med.*, **47**, 484 (1931).
5. Barach, A. L., and Richards, D. W., *Arch. Int. Med.*, **48**, 325 (1931).
6. Cournand, A., Brock, H. T., Rappaport, I., and Richards, D. W., *Arch. Int. Med.* (in press).

As long, therefore, as the breathing necessary for a given immediate metabolic state is accomplished easily, without undue sense of resistance in the air passages, or conscious mechanical effort referred to the ventilatory apparatus, there will be no dyspnea. Conversely, as is practically self-evident, whenever the action of the mechanics of pulmonary ventilation becomes difficult, as from resistance or obstruction in the air passages,

inelasticity of lung tissue, hypertonicity or asynergy of breathing muscles, or from increased work associated with large ventilatory need, then the increased proprioceptive impulses from these structures will induce a state of dyspnea. The dyspnea of physiological and of pathological conditions is thus essentially the same phenomenon. It is obvious that there may be, from one individual to another, or even in the same individual at different times, large differences in the extent of distress, or dyspnea, produced by a given mechanical state of the ventilatory apparatus; due to different psychological states, different degrees of "training" and other circumstances. This applies to normal as well as pathological subjects. There will also be large individual differences in the manner in which the metabolic requirements are fulfilled by the ventilatory apparatus; as illustrated by the greatly different levels of blood gases and volumes of ventilation given in the table. These questions are beyond the scope of the present brief note.

In general, clinical dyspnea is due to local difficulty or disturbance in the mechanical functioning of the organ of breathing (including in this term both the moving framework of the chest, and the lungs and air passages). Dyspnea is essentially independent of the physical chemistry of respiration considered as a function of the body as a whole.

The meaning of the above statement perhaps needs further definition. In states of oxygen want, for example, with arterial anoxemia, an existing dyspnea may often be quickly relieved by oxygen inhalation. By this procedure arterial anoxemia is reduced, there is correspondingly smaller respiratory stimulus, and pulmonary ventilation is less. It is our view that in this instance the actual relief of dyspnea is due to the easing of the mechanical effort of ventilation. The latter has of course been brought about by way of physicochemical readjustments.

The importance of mechanical disturbance in ventilation, as a cause of dyspnea, has been emphasized by various workers (Peabody,[6] Anthony,[7] Knipping[8] and others), but so far as we are aware, the general phenomenon of dyspnea has not previously been associated exclusively with sensory impulses from the chest and lungs, and the neuromuscular regulation of breathing; and separated from the physical chemistry of respiration of the body.

[1] Lumsden, T., "Observations on the Respiratory Centers," *Jour. Physiol.*, **57**, 153 354; **58**, 81, 111, 259 (1923).

[2] Hess, W. R., *Die Regulierung der Atmung*, Leipzig (1931).

[3] Heymans, C., Bouckarert, J. J., and Requiers, P., *La sinus carotidien et la zone homologue cardio-aortique*, Paris (1933).

[4] Gesell, R., and Moyer, C., "The Variability and Incidence of Types of Breathing in the Anaesthetized Dog," *Quart. Jour. Exp. Physiol.*, **24**, 315 (1935).

[5] Fleisch, A., "Neuere Ergebnisse über Mechanik und proprioceptive Bewegung der Atemsteuerung," *Ergebn. Physiol.*, **36**, 252 (1934).

[6] Peabody, F. W., and Wentworth, J. A., "Vital Capacity of the Lungs and Its Relation to Dyspnea," *Arch. Int. Med.*, **20**, 443 (1917).

[7] Anthony, A. J., "Zur Frage der Stenosenatmung," *Beitr. Kl. Tuberk.*, **70** (1929).

[8] Knipping, H. W., "Dyspnoe," *Ibid.*, **82**, 133 (1933).

ON TWO REMARKABLE CILIATE PROTOZOA FROM THE CAECUM OF THE INDIAN ELEPHANT

BY CHARLES A. KOFOID

DEPARTMENT OF ZOÖLOGY, UNIVERSITY OF CALIFORNIA

Read before the Academy, Monday, April 22, 1935

The contents of the digestive tract of herbivorous mammals afford an environment favorable to the establishment of a protozoan fauna. The food available for these Protozoa consists of finely comminuted fragments of vegetation, bits of cellulose, chromatophores, and other cellular contents, bacteria and other Protozoa supported by the food particles of the intestinal contents and by the fluid containing the products resulting from the action of the digestive enzymes. Additional factors favoring the invasion and maintenance of a protozoan fauna are controlled temperature, fairly constant food supply and an assured aqueous environment. A large volume of fluid and the long life of the host favor infection by the invading protozoan population and its continuous maintenance. The association of the young with the parents also assures the maintenance of a specific population in a specific host.

The organs inhabited by these Protozoa differ according to the feeding habits and the locations of the major digestive processes in the digestive tract of the host. Among ruminant animals the first and second stomachs are regions in which a highly diversified and exceedingly abundant protozoan fauna has evolved. In other mammals, such as the horse, a different ciliate fauna has been developed, not in the stomach, but in the caecum and colon. In the elephant a comparable development of another distinct ciliate fauna has occurred in the caecum and colon.

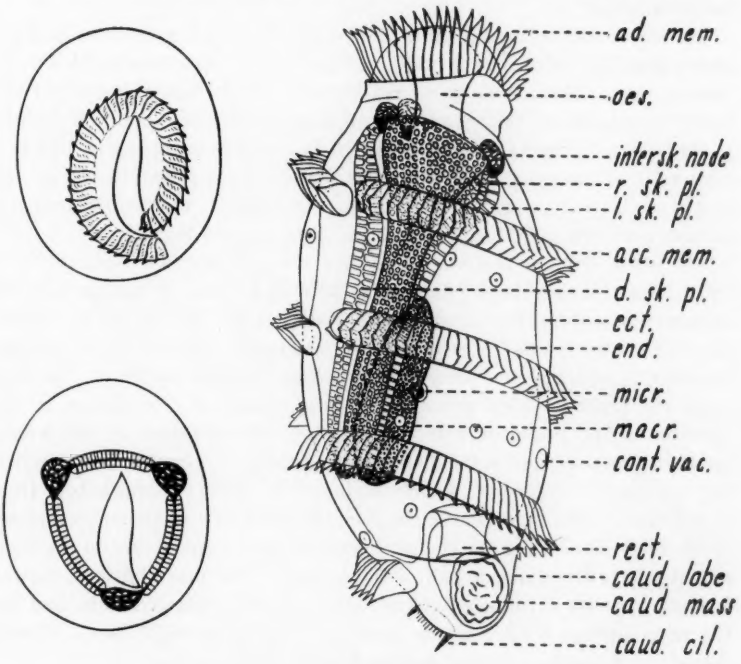
The purpose of this paper is to present two unique types of relatively very large and complicated ciliates found in the caecum and colon of the Indian elephant. The material for this investigation was secured in 1916 by special permission of the Maharajah of Mysore in the teak- and sandalwood forests in the Berambadi region on the slopes of the Nilgiri Mountains in South India. An elderly bull elephant, in the must, proscribed because of his destructive activities, was sacrificed to secure this material. The investigation was made possible by a grant from the Bache Fund of the National Academy of Sciences and from Regent Crocker of the University of California.

The two new genera of ciliates described in this paper belong in a new subfamily of the Ophryoscolecidae, a family hitherto confined to the ruminants and widely diversified into 17 genera and approximately 170 species. The Ophryoscolecidae form a compact but structurally diversified assemblage characterized by an adoral membranelle zone, a heavy peripheral pellicle, permanently located excretory vacuoles, and a system of from one to five vacuolated "skeletal" plates arranged peripherally to the pharynx and projecting posteriorly along the side of the macronucleus. These plates are characteristic of all but a few of the most primitive genera of this family. A few of the higher genera, such as *Epidinium*, *Diplodinium* and *Ophryoscolex*, of the Ophryoscolecidae, add to the adoral zone of membranelles a second, always dorsal, membranelle zone, having no relationship to the feeding function, but accessory to the locomotor function. All Ophryoscolecidae are powerful swimmers and ceaselessly bore their way through the finely divided material suspended in the fluid of the stomach. They are, therefore, heavy feeders and highly developed transformers of energy. The higher genera of this family have also developed in an orthogenetic fashion an increasing number of posterior spines which seem to have a storage function for some form of food reserve, a structural feature which would clearly impede reversal of the direction of locomotion. The excretory vacuoles, from one to twelve in number, are located in the dorsal side adjacent to the macronucleus, except in those cases where the vacuoles are numerous, in which instances they have a wider distribution.

The new subfamily of the Ophryoscolecidae, the Polydiniinae, found in the elephant, is characterized by the presence of numerous accessory membranelle zones, extending over the considerably elongated body. These zones, instead of being dorsal, are now divided bilaterally in two groups which still fall into a descending right spiral both individually and in pairs. There is thus a superposition of a secondary bilateral symmetry upon the primitive spiral one and an extension of metamerism by the added membranelle zones throughout the elongated body of these ciliates. The contractile vacuoles are greatly increased in number and arranged in zones parallel to the motor membranelle zones. This family is also characterized by the presence of the skeletal plates and has indications of a neuromotor system comparable to that of other Ophryoscolecidae.

Each of the two new genera is represented by a single species, the first, *Polydinium mysoreum*, and the second, *Elephantophilus zeta*. Polydinium is the smaller (200–250μ) of the two. It is characterized by five initial additional membranelle zones, increasing to seven during early stages of growth, and ultimately to ten, prior to binary fission. This genus also has three equal club-shaped, double-layered, skeletal plates running nearly the whole length of the body surrounding the oesophagus. The plates are crowned by the oesophageal neuromotor ring with deeply stain-

ing enlargements at the nodes. There are from twenty to thirty excretory vacuoles distributed in irregular rows posterior to each of the pairs of



Polydinium mysoreum

FIGURES 1 TO 3

Figure 1. *Polydinium mysoreum* gen. nov., sp. nov., from *Elephas inducus* Cuvier. View of right side. × 550.

Abr.: *acc. mem.*, accessory membranelle zone; *ad. mem.*, adoral membranelle; *caud. cil.*, caudal cilia; *caud. lobe*, caudal lobe; *caud. mass.*, caudal mass; *cont. vac.*, contractile vacuole; *d. sk. pl.*, dorsal skeletal plate; *ect.*, ectoplasm; *end.*, endoplasm; *intersk. node*, interskeletal node of oesophageal neural ring; *l. sk. pl.*, left skeletal plate; *macr.*, macronucleus; *micr.* micronuleus; *oes.*, oesophagus; *r. sk. pl.*, right skeletal plate; *rect.*, rectum.

Figure 2. Cross-section of same, showing oesophagus and adoral membranelle zone. × 550.

Figure 3. Oblique section through anterior ends of the skeletal plates showing their bilaminate structure, circumoesophageal neural ring, and two dorsal and one ventral interskeletal nodes. × 550.

accessory membranelle zones. The body has a posterior lobe with an accessory group of cilia on its dorsal face. This lobe always contains an

irregular mass of stored material of unknown composition in a vacuole in the ectoplasm. It lies behind the anal opening and is apparently not faecal material.

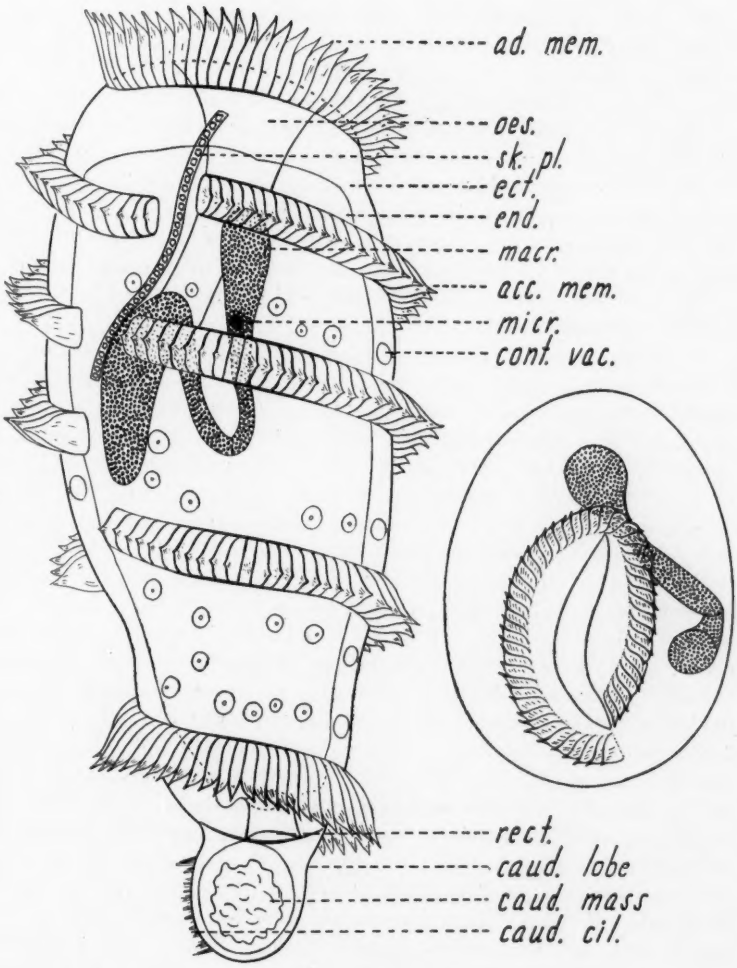
The macronucleus is stout club-shaped, lies on the left side of the skeletal plates near the middle of the body, and has the micronucleus bedded in its ventral side. The adoral spiral makes but a single turn and is relatively feebly developed in this genus as compared with other ophryoscolecids. It leads directly into the well-developed but narrow oesophagus, which is bilaterally compressed and continues posteriorly for two-thirds of the length of the body, opening there in the endoplasm. The animal feeds on minute particles, including plant debris, bacteria and flagellates.

The other genus is represented by but a single species, *Elephantophilus zeta*. This is a very large ciliate, 250–290μ in length. It has six seriated accessory membranelle zones in the young stage, increasing to twelve prior to binary fission. These also are seriated in bilateral pairs in a descending right spiral and present a secondary bilateral symmetry imposed upon the primary spiral symmetry. It is similar to *Polydinium* in the position of the posterior lobe, the relative subordination of the adoral membranelle zone, and in the distribution of the excretory vacuoles paralleling successive accessory membranelles. It differs remarkably from *Polydinium* in the reduction of the skeletal plates to a single one composed of one linear row of prismatic chambers forming a sigmoid line on the right dorsal side in the anterior half of the animal. The macronucleus, viewed laterally, has the form of a flattened letter *Z*, with enlarged ends, and has the micronucleus bedded in the posterior lobe on its right face. Viewed dorsally it presents a slightly flattened spiral form.

The oesophagus is much wider and shorter than that of *Polydinium* and lacks the skeletal support. The food of this organism often includes a large fragment of woody tissue which in some cases extends beyond both ends of the body, distorting the animal which is passing this food fragment through its endoplasm.

At the time of binary fission in both genera the number of accessory membranelle zones is doubled, and new excretory vacuoles arise in connection with each incipient bilateral addition to these zones. The zones are not added at one end of the body, but are interpolated between the initial set of zones on alternate sides. These added zones develop progressively from the anterior end posteriorly and the period of their development is evidently prolonged, because of the fact that all stages are approximately equally distributed in number among the individuals found in the stomach. There appears to be no predominant stage in which the individual persists for a longer period.

Discussion.—Viewed from the standpoint of the evolution of the Ophryoscolecidae, the family Polydiniinae represents the highest orthogenetic



FIGURES 4 AND 5

Figure 4. *Elephantophilus zeta* gen. nov., sp. nov., from *Elephas indicus* Cuvier. View of right side. × 550. Abbreviations as in figure 1.

Figure 5. Cross-section of same, showing oesophagus with optical projections of adoral membranelle zone and macronucleus. × 550.

phase of the ophryoscolecids. The organelle systems sharing in this evolution are primarily the locomotor and excretory systems, indicating a basic functional evolution of locomotor activities. This evolution in these ciliate Protozoa exhibits not only an orthogenetic aspect, but it involves the principle of metamerism, the repetition along the major axis of motor and excretory organelles. This is also a primary feature of the evolution of coelomate Metazoa in which muscle masses, bilateral appendages and excretory organelles have a metameric sequence. In the ophryoscolecids with one dorsal membranelle zone a fibril of the neuromotor system accompanies the membranelle zone. It is to be expected that the repeated membranelles of this subfamily will exhibit a repetition of the neuromotor fibrils.

Superposed upon the primary spiral structure of these ciliates there is an incipient bilateral symmetry shown by the flattening of the oesophagus in the sagittal plane and the division of the added membranelles, originally introduced as a single dorsal one, into a right and left segment in each metamere. This bilaterality is still subject to the spirality of the total accessory membranelle equipment. The skeletal plates and the nodes, or enlargements, on the oesophageal ring of the neuromotor system both show a dorsoventral differentiation, there being two nodes and one middorsal plate on the dorsal side, and two lateral plates, and one ventral node on the oesophageal ring, on the ventral side.

The Proboscidea form one of the outstanding achievements of the evolutionary process among the higher Mammalia. Their protozoan parasites, or commensals, have presumably accompanied their hosts during the process of their evolution. We find in *Polydinium* and *Elephantophilus* the highest development known in the family Ophryoscolecidae. The evolution of the host and that of its parasites are coincident processes and are comparable in direction. They are alike in that each culminates in greater size, greater complexity and greater diversification of the characteristic organs which distinguish the respective lines of the evolving orthogenetic series. The basic principles of antero-posterior differentiation, metameric repetition, bilateral symmetry, correlation of locomotor, excretory and nervous systems, appear in these Protozoa in a manner analogous to their appearance in the metameric coelomate Metazoa.